

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

PCT

ORGANISATION MONDIALE DE LA PROPRIÉTÉ INTELLECTUELLE  
Bureau international

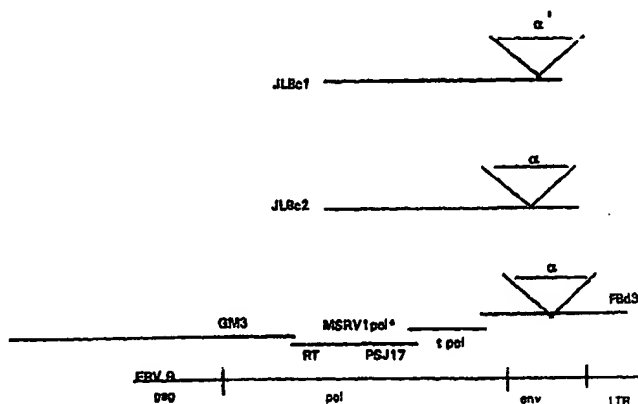


DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIERE DE BREVETS (PCT)

<p>(51) Classification internationale des brevets <sup>6</sup> : <b>C12N 15/48, 7/00, C07K 14/15, C12Q 1/68, C07K 16/10, G01N 33/569, A61K 39/21, 39/42, 48/00</b></p>	<p><b>A1</b></p>	<p>(11) Numéro de publication internationale: <b>WO 97/06260</b> (43) Date de publication internationale: 20 février 1997 (20.02.97)</p>
<p>(21) Numéro de la demande internationale: PCT/FR96/01244 (22) Date de dépôt international: 2 août 1996 (02.08.96) (30) Données relatives à la priorité: 95/09643 3 août 1995 (03.08.95) FR (71) Déposant (pour tous les Etats désignés sauf US): BIO MERIEUX [FR/FR]; F-69280 Marcy-l'Etoile (FR). (72) Inventeurs; et (75) Inventeurs/Déposants (US seulement): PERRON, Hervé [FR/FR]; 134, rue du Docteur-E.-Locard, F-69005 Lyon (FR). BESEME, Frederic [FR/FR]; 39, rue de la Noyera, F-38090 Villefontaine (FR). BEDIN, Frédéric [FR/FR]; 6, rue Gaspard-André, F-69002 Lyon (FR). PARANHOS-BACCALA, Glaucia [FR/FR]; 75, cours Gambetta, F-69003 Lyon (FR). KOMURIAN-PRADEL, Florence [FR/FR]; Chemin Vial, F-69450 Saint-Cyr-au-Mont-d'Or (FR). JOLIVET-REYNAUD, Colette [FR/FR]; 16, avenue des Colonnes, F-69500 Bron (FR). MANDRAND, Bernard [FR/FR]; 21, rue de la Doua, F-69100 Villeurbanne (FR). (74) Mandataire: CABINET GERMAIN &amp; MAUREAU; Boîte postale 3011, F-69392 Lyon Cédex 03 (FR).</p>	<p>(81) Etats désignés: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, brevet ARIPO (KE, LS, MW, SD, SZ, UG), brevet eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Publiée</b> Avec rapport de recherche internationale. Avec une indication relative à un micro-organisme déposé, fournie selon la règle 13 bis, séparément, et non pas avec la description. Date de réception par le bureau international: 23 septembre 1996 (23.09.96)</p>	

(54) Title: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS ASSOCIATED WITH MULTIPLE SCLEROSIS USEFUL FOR DIAGNOSTIC, PREVENTIVE AND THERAPEUTIC PURPOSES

(54) Titre: MATERIEL VIRAL ET FRAGMENTS NUCLEOTIDIQUES ASSOCIES A LA SCLEROSE EN PLAQUES, A DES FINS DE DIAGNOSTIC, PROPHYLACTIQUES ET THERAPEUTIQUES



(57) Abstract

A viral material, in isolated or purified form, having a genome which includes a nucleotide sequence selected from the group of sequences SEQ ID NO 46, SEQ ID NO 51, SEQ ID NO 52, SEQ ID NO 53 and SEQ ID NO 56, their complementary sequences and equivalent sequences, particularly the nucleotide sequences having, for any chain of 100 adjacent monomers, a respective homology of 50 % or more and preferably 70 % or more with SEQ ID NO 46, SEQ ID NO 51, SEQ ID NO 52, SEQ ID NO 53 and SEQ ID NO 56, and the complementary sequences thereof.

**(57) Abrégé**

Matériel viral, à l'état isolé ou purifié, dont le génome comprend une séquence nucléotidique choisie dans le groupe incluant les séquences SEQ ID NO 46, SEQ ID NO 51, SEQ ID NO 52, SEQ ID NO 53 et SEQ ID NO 56, leurs séquences complémentaires, et leurs séquences équivalentes, notamment les séquences nucléotidiques présentant, pour toute suite de 100 monomères contigus, au moins 50 % et préférentiellement au moins 70 % d'homologie avec respectivement lesdites séquences SEQ ID NO 46, SEQ ID NO 51, SEQ ID NO 52, SEQ ID NO 53 et SEQ ID NO 56, et leurs séquences complémentaires.

**UNIQUEMENT A TITRE D'INFORMATION**

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AT	Arménie	GB	Royaume-Uni	MW	Malawi
AT	Autriche	GE	Géorgie	MX	Mexique
AU	Australie	GN	Guinée	NE	Niger
BB	Barbade	GR	Grèce	NL	Pays-Bas
BE	Belgique	HU	Hongrie	NO	Norvège
BF	Burkina Faso	IE	Irlande	NZ	Nouvelle-Zélande
BG	Bulgarie	IT	Italie	PL	Pologne
BJ	Bénin	JP	Japon	PT	Portugal
BR	Brsil	KE	Kenya	RO	Roumanie
BY	Bélarus	KG	Kirghizistan	RU	Fédération de Russie
CA	Canada	KP	République populaire démocratique de Corée	SD	Soudan
CF	République centrafricaine	KR	République de Corée	SE	Suède
CG	Congo	KZ	Kazakhstan	SG	Singapour
CH	Suisse	LI	Liechtenstein	SI	Slovénie
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovaquie
CM	Cameroun	LR	Libéria	SN	Sénégal
CN	Chine	LT	Lituanie	SZ	Swaziland
CS	Tchécoslovaquie	LU	Luxembourg	TD	Tchad
CZ	République tchèque	LV	Lettonie	TG	Togo
DE	Allemagne	MC	Monaco	TJ	Tadjikistan
DK	Danemark	MD	République de Moldova	TT	Trinité-et-Tobago
EE	Estonie	MG	Madagascar	UA	Ukraine
ES	Espagne	ML	Mali	UG	Ouganda
FI	Finlande	MN	Mongolie	US	Etats-Unis d'Amérique
FR	France	MR	Mauritanie	UZ	Ouzbékistan
GA	Gabon			VN	Viet Nam

**MATERIEL VIRAL ET FRAGMENTS NUCLEOTIDIQUES ASSOCIES A LA  
SCLEROSE EN PLAQUES, A DES FINS DE DIAGNOSTIC,  
PROPHYLACTIQUES ET THERAPEUTIQUES**

5           La Sclérose en Plaques (SEP) est une maladie  
démýélinisante du système nerveux central (SNC) dont la  
cause complète reste encore inconnue.

De nombreux travaux ont étayé l'hypothèse d'une  
étiologie virale de la maladie, mais aucun des virus  
10 connus testés ne s'est avéré être l'agent causal  
recherché: une revue des virus recherchés depuis des  
années dans la SEP a été faite par E. Norrby (1) et R.T.  
Johnson (2).

Récemment, un rétrovirus, différent des  
15 rétrovirus humains connus a été isolé chez des patients  
atteints de SEP (3,4 et 5). Les auteurs ont aussi pu  
montrer que ce rétrovirus pouvait être transmis in vitro,  
que des patients atteints de SEP produisaient des  
anticorps susceptibles de reconnaître des protéines  
20 associées à l'infection des cellules leptoméningées par ce  
rétrovirus, et que l'expression de ce dernier pouvait être  
fortement stimulée par les gènes immédiats-précoces de  
certains herpesvirus (6).

Tous ces résultats plaident en faveur du rôle  
25 dans la SEP d'au moins un rétrovirus inconnu ou d'un virus  
ayant une activité transcriptase inverse détectable selon  
la méthode publiée par H. Perron (3) et qualifiée  
d'activité "RT de type LM7". Le contenu de la publication  
identifiée par (3) est incorporé à la présente  
30 description, par référence.

Récemment, les travaux de la Demanderesse ont  
permis d'obtenir deux lignées continues de cellules  
infectées par des isolats naturels provenant de deux  
patients différents atteints de SEP, par un procédé de  
35 culture tel que décrit dans le document WO-A-93 20188,  
dont le contenu est incorporé par référence à la présente



description. Ces deux lignées dérivées de cellules de plexus-choroïdes humains, dénommées LM7PC et PLI-2 ont été déposées à l'E.C.A.C.C. respectivement le 22 juillet 1992 et le 8 janvier 1993, sous les numéros 92 072201 et 5 93 010817, conformément aux dispositions du Traité de Budapest. Par ailleurs, les isolats viraux possédant une activité RT de type LM7 ont également été déposés à l'E.C.A.C.C. sous la dénomination globale de "souches". La "souche" ou isolat hébergé par la lignée PLI-2, dénommée 10 POL-2, a été déposée auprès de l'E.C.A.C.C. le 22 juillet 1992 sous le n° V92072202. La "souche" ou isolat hébergé par la lignée LM7PC, dénommée MS7PG, a été déposée auprès de l'E.C.A.C.C. le 8 janvier 1993 sous le n° V93010816.

15 A partir des cultures et des isolats précités, caractérisés par des critères biologiques et morphologiques, on s'est ensuite attaché à caractériser le matériel nucléaire associé aux particules virales produites dans ces cultures.

20 Les portions de génome déjà caractérisées ont été utilisées pour mettre au point des tests de détection moléculaire du génome viral et des tests immunosérologiques, utilisant les séquences d'acides aminés codés par les séquences nucléotidiques du génome viral, 25 pour détecter la réponse immunitaire dirigée contre des épitopes associés à l'infection et/ou l'expression virale.

Ces outils ont déjà permis de confirmer une association entre la SEP et l'expression des séquences identifiées dans les brevets cités plus loin. Cependant, 30 le système viral découvert par la Demanderesse, s'apparente à un système rétroviral complexe. En effet, les séquences retrouvées encapsidées dans les particules virales extracellulaires produites par les différentes cultures de cellules de patients atteints de SEP, montrent 35 clairement qu'il y a co-encapsidation de génomes rétroviraux apparentés, mais différents du génome

rétroviral "sauvage" qui produit les particules virales infectantes. Ce phénomène a été observé entre des rétrovirus réplicatifs et des rétrovirus endogènes appartenant à la même famille, voire même hétérologues. La notion de rétrovirus endogène est très importante dans le contexte de notre découverte car, dans le cas de MSRV-1, on a observé que des séquences rétrovirales endogènes comprenant des séquences homologues au génome MSRV-1, existent dans l'ADN humain normal. L'existence d'éléments rétroviraux endogènes (ERV) apparentés à MSRV-1 par tout ou partie de leur génome, explique le fait que l'expression du rétrovirus MSRV-1 dans les cellules humaines puisse interagir avec des séquences endogènes proches. Ces interactions sont retrouvées dans le cas de rétrovirus endogènes pathogènes et/ou infectieux (par exemple certaines souches écotropes du Murine Leukaemia virus), dans le cas de rétrovirus exogènes dont la séquence nucléotidique peut être retrouvée partiellement ou en totalité, sous forme d'ERVs, dans le génome de l'animal hôte (ex. virus exogène de la tumeur mammaire de la souris transmis par le lait). Ces interactions consistent principalement en (i) une transactivation ou co-activation d'ERVs par le rétrovirus réplicatif, (ii) une encapsidation "illégitime" d'ARN apparentés d'ERVs, ou d'ERVs -voire d'ARN cellulaires- possédant simplement des séquences d'encapsidation compatibles, dans les particules rétrovirales produites par l'expression de la souche répllicative, parfois transmissibles et parfois avec une pathogénicité propre, et (iii) des recombinaisons plus ou moins importantes entre les génomes co-encapsidés, notamment dans les phases de transcription inverse, qui conduisent à la formation de génomes hybrides, parfois transmissibles et parfois avec une pathogénicité propre.

Ainsi, (i) différentes séquences apparentées à MSRV-1 ont été retrouvées dans les particules virales purifiées; (ii) l'analyse moléculaire des différentes

régions du génome rétroviral MSRV-1 doit être faite en analysant systématiquement les séquences co-encapsidées, interférantes et/ou recombinées qui sont générées par l'infection et/ou l'expression de MSRV-1, de plus, certains clones peuvent avoir des parties de séquences défectives produites par la réplication rétrovirale et les erreurs de matrice et/ou de transcription de la transcriptase inverse; (iii) les familles de séquences apparentées à une même région génomique rétrovirale sont les supports d'une détection diagnostique globale qui peut être optimisée par l'identification de régions invariables parmi les clones exprimés et par l'identification de trames de lectures responsables de la production de polypeptides antigéniques et/ou pathogènes qui peuvent n'être produits que par une partie, voire un seul, des clones exprimés et dans ces conditions, l'analyse systématique des clones exprimés dans une région d'un gène donné permet d'évaluer la fréquence de variation et/ou de recombinaison du génome MSRV-1 dans cette région et de définir les séquences optimales pour les applications, notamment diagnostiques; (iv) la pathologie provoquée par un rétrovirus tel que MSRV-1 peut être un effet direct de son expression et des protéines ou peptides produits de ce fait, mais aussi un effet de l'activation, de l'encapsidation, de la recombinaison de génomes apparentés ou hétérologues et des protéines ou peptides produits de ces faits ; ainsi ces génomes associés à l'expression de et/ou l'infection par MSRV-1 sont-ils une partie intégrante de la pathogénicité potentielle de ce virus et donc, constituent des supports de détection diagnostique et des cibles thérapeutiques particulières. De même tout agent associé à, ou, co-facteur de ces interactions responsables de la pathogénie en cause, tel que MSRV-2 ou le facteur gliotoxique décrits dans la demande de brevet publiée sous le N° FR-2 716 198, peut participer à l'élaboration d'une stratégie globale et très efficace de

diagnostic, de pronostic, de suivi thérapeutique et/ou de thérapeutique intégrée de la SEP notamment, mais aussi de toute autre maladie associée aux mêmes agents.

Dans ce contexte, on a fait une découverte  
5 parallèle dans une autre maladie autoimmune, la polyarthrite rhumatoïde (PR), qui a été décrite dans la demande de brevet français déposée sous le N°95 02960. Cette découverte montre que, en appliquant des approches méthodologiques similaires à celles qui furent utilisées  
10 dans les travaux de la Demanderesse sur la SEP, on a pu identifier un rétrovirus exprimé dans la PR qui partage les séquences décrites pour MSRV-1 dans la SEP et aussi, la co-existence d'une séquence associée MSRV-2 également décrite dans la SEP. En ce qui concerne MSRV-1, les  
15 séquences détectées communément dans la SEP et la PR, concernent les gènes *pol* et *gag*. En l'état actuel des connaissances, on peut associer les séquences *gag* et *pol* décrites aux souches MSRV-1 exprimées dans ces deux maladies.

20 La présente demande de brevet a pour objet différents résultats, supplémentaires par rapport à ceux déjà protégés par les demandes de brevet français :

- N° 92 04322 du 03.04.1992, publiée sous le N° 2 689 519 ;
  - N° 92 13447 du 03.11.1992, publiée sous le N° 2 689 521 ;
  - 25 -N° 92 13443 du 03.11.1992, publiée sous le N° 2 689 520 ;
  - N° 94 01529 du 04.02.1994, publiée sous le N° 2 715 936 ;
  - N° 94 01531 du 04.02.1994, publiée sous le N° 2 715939 ;
  - N° 94 01530 du 04.02.1994, publiée sous le N° 2 715 936 ;
  - N° 94 01532 du 04.02.1994, publiée sous le N° 2 715 937 ;
  - 30 -N° 94 14322 du 24.11.1994, publiée sous le N° 2 727 428 ;
- et
- N° 94 15810 du 23.12.1994, publiée sous le N° 2 728 585.

La présente invention concerne tout d'abord un matériel viral, à l'état isolé ou purifié, pouvant être  
35 appréhendé ou caractérisé de différentes manières :

- son génome comprend une séquence nucléotidique choisie dans le groupe incluant les séquences SEQ ID NO 46, SEQ ID NO 51, SEQ ID NO 52, SEQ ID NO 53, SEQ ID NO 56, SEQ ID NO 58, SEQ ID NO 59, SEQ ID NO 60, SEQ ID NO 61, SEQ ID NO 89, leurs séquences complémentaires, et leurs séquences équivalentes, notamment les séquences nucléotidiques présentant, pour toute suite de 100 monomères contigus, au moins 50 % et préférentiellement au moins 70 % d'homologie avec respectivement lesdites séquences SEQ ID NO 46, SEQ ID NO 51, SEQ ID NO 52, SEQ ID NO 53, SEQ ID NO 56, SEQ ID NO 58, SEQ ID NO 59, SEQ ID NO 60, SEQ ID NO 61, SEQ ID NO 89, et leurs séquences complémentaires ;
- la région de son génome comprenant les gènes *env*, *pol* et une partie du gène *gag*, à l'exclusion de la sous-région ayant une séquence identique ou équivalente à SEQ ID NO 1, code pour tout polypeptide présentant, pour toute suite contigue d'au moins 30 acides aminés, au moins 50 %, et de préférence au moins 70 % d'homologie avec une séquence peptidique codée par toute séquence nucléotidique choisie dans le groupe incluant SEQ ID NO 46, SEQ ID NO 51, SEQ ID NO 52, SEQ ID NO 53, SEQ ID NO 56, SEQ ID NO 58, SEQ ID NO 59, SEQ ID NO 60, SEQ ID NO 61, SEQ ID NO 89, et leurs séquences complémentaires ;
- le gène *pol* comprend une séquence nucléotidique identique ou équivalente, partiellement ou totalement, à SEQ ID NO 57, à l'exclusion de SEQ ID NO 1.
- le gène *gag* comprend une séquence nucléotidique identique ou équivalente, partiellement ou totalement, à SEQ ID NO 88.

Comme évoqué précédemment, selon la présente invention, le matériel viral tel que défini précédemment est associé à la SEP. Et tel que défini par référence au gène *pol* ou *gag* de MSRV-1, et plus particulièrement aux

séquences SEQ ID NOS 51, 56, 57, 59, 60, 61, 88 et 89, ce matériel viral est associé à la PR.

La présente invention concerne également différents fragments nucléotidiques, comprenant chacun une  
5 séquence nucléotidique choisie dans le groupe incluant :

- (a) toutes les séquences génomiques, partielles et totales, du gène *pol* du virus MSRV-1, sauf la séquence totale du fragment nucléotidique défini par SEQ ID NO 1 ;
- 10 (b) toutes les séquences génomiques, partielles et totales, du gène *env* de MSRV-1 ;
- (c) toutes les séquences génomiques partielles du gène *gag* de MSRV-1 ;
- (d) toutes les séquences génomiques chevauchant le gène  
15 *pol* et le gène *env* du virus MSRV-1, et chevauchant le gène *pol* et le gène *gag* ;
- (e) toutes les séquences, partielles et totales, d'un clone choisi dans le groupe incluant les clones FBd3 (SEQ ID NO 46), t *pol* (SEQ ID NO 51), JLbc1 (SEQ ID NO 52), JLbc2 (SEQ ID NO 53), GM3 (SEQ ID NO 56), FBd13 (SEQ ID NO 58), LB19 (SEQ ID NO 59), LTRGAG12 (SEQ ID NO 60), FP6 (SEQ ID NO 61), G+E+A (SEQ ID NO 89), à l'exclusion de  
20 toute séquence nucléotidique identique à ou comprise dans la séquence définie par SEQ ID NO 1 ;
- 25 (f) les séquences complémentaires auxdites séquences génomiques ;
- (g) les séquences équivalentes auxdites séquences (a) à (e), notamment les séquences nucléotidiques  
30 présentant, pour toute suite de 100 monomères contigus, au moins 50 % et de préférence au moins 70 % d'homologie avec lesdites séquences (a) à (d),  
à condition que ce fragment nucléotidique ne comprenne pas ou ne consiste pas en la séquence ERV-9 telle que décrite  
35 dans LA MANTIA et al. (18).

Par séquences génomiques, partielles ou totales, on inclut toutes séquences associées par co-encapsidation ou par co-expression, ou recombinées.

Préférentiellement, un tel fragment comprend :

- 5 - ou une séquence nucléotidique identique à une séquence génomique partielle ou totale du gène *pol* du virus MSRV-1, sauf à la séquence totale du fragment nucléotidique défini par SEQ ID NO 1, ou identique à toute séquence équivalente à ladite séquence génomique  
10 partielle ou totale, notamment homologue à cette dernière ;
- ou une séquence nucléotidique identique à une séquence génomique partielle ou totale du gène *env* du virus MSRV-1, ou identique à toute séquence complémentaire de  
15 ladite séquence nucléotidique, ou identique à toute séquence équivalente à ladite séquence nucléotidique, notamment homologue à cette dernière.

De manière particulière, l'invention concerne un fragment nucléotidique comprenant une séquence  
20 nucléotidique codante, partiellement ou totalement identique à une séquence nucléotidique choisie dans le groupe incluant :

- la séquence nucléotidique définie par SEQ ID NO 40, SEQ ID NO 62 ou SEQ ID NO 89 ;
- 25 -les séquences complémentaires à SEQ ID NO 40, SEQ ID NO 62 ou SEQ ID NO 89 ;
- les séquences équivalentes et notamment homologues à SEQ ID NO 40, SEQ ID NO 62 ou SEQ ID NO 89 ;
- les séquences codant pour tout ou partie de la séquence  
30 peptidique définie par SEQ ID NO 39, SEQ ID NO 63 ou SEQ ID NO 90 ;
- les séquences codant pour tout ou partie d'une séquence peptidique équivalente, notamment homologue à  
35 SEQ ID NO 39, SEQ ID NO 63 ou SEQ ID NO 90, susceptible d'être reconnue par des séra de patients infectés par le

virus MSRV-1, ou chez lesquels le virus MSRV-1 a été réactivé.

L'invention concerne aussi toute sonde nucléique de détection d'un agent pathogène et/ou infectant associé à la SEP, susceptible de s'hybrider spécifiquement sur tout fragment tel que précédemment défini, appartenant ou compris dans le génome dudit agent pathogène. Elle concerne en outre toute sonde nucléique de détection d'un agent pathogène et/ou infectant associé à la PR, susceptible de s'hybrider spécifiquement sur tout fragment tel que défini précédemment par référence aux gènes *pol* et *gag*, et en particulier par rapport aux séquences SEQ ID NOS 40, 51, 56, 59, 60, 61, 62, 89 et SEQ ID NOS 39, 63 et 90.

L'invention concerne aussi une amorce pour l'amplification par polymérisation d'un ARN ou d'un ADN d'un matériel viral, comprenant une séquence nucléotidique identique ou équivalente à au moins une partie de la séquence nucléotidique de tout fragment tel que défini précédemment, notamment une séquence nucléotidique présentant pour toute suite de 10 monomères contigus, au moins 70 % d'homologie avec au moins ladite partie dudit fragment. Préférentiellement, la séquence nucléotidique d'une telle amorce est identique à l'une quelconque des séquences choisies dans le groupe incluant SEQ ID NO 47 à SEQ ID NO 50, SEQ ID NO 55 et SEQ ID NO 64 SEQ ID NO 86.

De manière générale, l'invention embrasse également tout ARN ou ADN, et notamment vecteur de répllication, comprenant un fragment génomique du matériel viral tel que précédemment défini, ou un fragment nucléotidique tel que précédemment défini.

L'invention s'intéresse également aux différents peptides codés par tout cadre de lecture ouvert appartenant à un fragment nucléotidique tel que précédemment défini, notamment tout polypeptide, par exemple tout oligopeptide formant ou comprenant un



déterminant antigénique reconnu par des sera de patients infectés par le virus MSRV-1, et/ou chez lesquels le virus MSRV-1 a été réactivé. Préférentiellement, ce polypeptide est antigénique, et est codé par le cadre de lecture  
5 ouvert commençant dans le sens 5'-3', au nucléotide 181, et finissant au nucléotide 330 de SEQ ID NO 1.

A titre particulier, l'invention concerne un polypeptide antigénique, reconnu par les sera de patients infectés par le virus MSRV-1, et/ou chez lesquels le virus  
10 MSRV-1 a été réactivé, dont la séquence peptidique est identique, partiellement ou totalement, ou équivalente à la séquence définie par SEQ ID NO 39, SEQ ID NO 63 et SEQ ID NO 87 ; une telle séquence est identique par exemple à toute séquence choisie dans le groupe incluant  
15 les séquences SEQ ID NO 41 à SEQ ID NO 44, SEQ ID NO 63 et SEQ ID NO 87.

La présente invention propose également des anticorps mono- ou polyclonaux, dirigés contre le virus MSRV-1, obtenus par réaction immunologique d'un  
20 organisme humain ou animal, à un agent immunogène constitué par un polypeptide antigénique tel que défini précédemment.

L'invention s'intéresse ensuite :

- aux réactifs de détection du virus MSRV-1, ou d'une  
25 exposition à ce dernier, comprenant à titre de substance réactive, un peptide, notamment antigénique, tel que précédemment défini, ou un anti-ligand, notamment anticorps dudit peptide ;
- à toutes compositions diagnostiques, prophylactiques, ou  
30 thérapeutiques, comprenant un ou plusieurs peptides, notamment antigéniques, tels que précédemment définis, ou un ou plusieurs anti-ligands, notamment anticorps des peptides, considérés précédemment ; une telle composition est préférentiellement, et à titre  
35 d'exemple, une composition vaccinale.

L'invention s'intéresse également à toute composition diagnostique, prophylactique, ou thérapeutique, notamment pour inhiber l'expression d'au moins un agent pathogène et/ou infectant associé à la SEP, 5 comprenant un fragment nucléotidique tel que précédemment défini, ou un polynucléotide, notamment oligonucléotide, dont la séquence est partiellement identique à celle dudit fragment, sauf à celle du fragment ayant la séquence nucléotidique SEQ ID NO 1. De la même manière, elle 10 s'intéresse à toute composition diagnostique, prophylactique, ou thérapeutique, notamment pour inhiber l'expression d'au moins un agent pathogène et/ou infectant associé à la PR, comprenant un fragment nucléotidique tel que précédemment défini par référence aux gènes *pol* et 15 *gag*, et en particulier par rapport aux séquences SEQ ID NOS 40, 51, 56, 59, 60, 61, 62 et 89 .

Selon l'invention, ces mêmes fragments, ou polynucléotides, notamment oligonucléotides, peuvent entrer dans toutes compositions appropriées, pour 20 détecter, selon tout procédé ou méthode convenable, un agent pathologique, et/ou infectant, associé respectivement à la SEP et à la PR, dans un échantillon biologique. Dans un tel procédé, on met en contact un ARN et/ou un ADN présumé appartenir ou provenant dudit agent 25 pathologique et/ou infectant, et/ou leur ARN et/ou ADN complémentaire, avec une telle composition.

La présente invention concerne également tout procédé pour détecter la présence ou l'exposition à un tel agent pathologique et/ou infectant, dans un échantillon 30 biologique, en mettant en contact cet échantillon avec un peptide, notamment antigénique, tel que précédemment défini, ou un anti-ligand, notamment anticorps de ce peptide, tel que précédemment défini.

En pratique, et par exemple, un dispositif de 35 détection du virus MSRV-1 comprend un réactif tel que précédemment défini, supporté par un support solide,

immunologiquement compatible avec le réactif, et un moyen de mise en contact de l'échantillon biologique, par exemple un échantillon de sang ou de liquide céphalo-rachidien, susceptible de contenir des anticorps anti-MSRV-1, avec ce réactif, dans des conditions permettant une éventuelle réaction immunologique, tout ceci avec des moyens de détection du complexe immun formé avec ce réactif.

Pour terminer, l'invention concerne également la détection d'anticorps anti-MSRV-1, dans un échantillon biologique, par exemple un échantillon de sang ou de liquide céphalo-rachidien, selon lequel on met en contact cet échantillon avec un réactif tel que précédemment défini, consistant en un anticorps, dans des conditions permettant leur éventuelle réaction immunologique, et on détecte ensuite la présence du complexe immun ainsi formé avec le réactif.

Avant de détailler l'invention, différents termes utilisés dans la description et les revendications sont à présent définis :

- par souche ou isolat, on entend toute fraction biologique infectante et/ou pathogène, contenant par exemple des virus et/ou des bactéries et/ou des parasites, générant un pouvoir pathogène et/ou antigénique, hébergée par une culture ou un hôte vivant ; à titre d'exemple, une souche virale selon la définition précédente peut contenir un agent co-infectant, par exemple un protiste pathogène,

- le terme "MSRV" utilisé dans la présente description désigne tout agent pathogène et/ou infectant, associé à la SEP, notamment une espèce virale, les souches atténuées de ladite espèce virale, ou les particules défectives interférentes ou contenant des génomes co-encapsidés ou encore des génomes recombinés avec une partie du génome MSRV-1, dérivées de cette espèce. Il est connu que les virus et particulièrement les virus contenant de l'ARN ont une variabilité, consécutive

notamment à des taux relativement élevés de mutation spontanée (7), dont il sera tenu compte ci-après pour définir la notion d'équivalence,

- par virus humain, on entend un virus  
5 susceptible d'infecter ou d'être hébergé par l'être humain,

- compte tenu de toutes les variations et/ou recombinaison naturelles ou induites, pouvant être rencontrées dans la pratique de la présente invention, les  
10 objets de cette dernière, définis ci-dessus et dans les revendications, ont été exprimés en comprenant les équivalents ou dérivés des différents matériels biologiques définis ci-après, notamment des séquences homologues nucléotidiques ou peptidiques,

15 - le variant d'un virus ou d'un agent pathogène et/ou infectant selon l'invention, comprend au moins un antigène reconnu par au moins un anticorps dirigé contre au moins un antigène correspondant dudit virus et/ou dudit agent pathogène et/ou infectant, et/ou un génome dont  
20 toute partie est détectée par au moins une sonde d'hybridation, et/ou au moins une amorce d'amplification nucléotidique spécifique dudit virus et/ou agent pathogène et/ou infectant, comme par exemple pour le virus dit MSRV-1, celles ayant une séquence nucléotidique choisie  
25 parmi SEQ ID N° 20 à SEQ ID N° 24, SEQ ID N° 26, SEQ ID N° 16 à SEQ ID N° 19, SEQ ID N° 31 à SEQ ID N° 33, SEQ ID N° 45, SEQ ID N° 47, SEQ ID N° 48, SEQ ID N° 49, SEQ ID N° 50, SEQ ID N° 45 et leurs séquences complémentaires, dans des conditions d'hybridation  
30 déterminées bien connues de l'homme de l'art,

- selon l'invention, un fragment nucléotidique ou un oligonucléotide ou un polynucléotide est un enchaînement de monomères, ou un biopolymère, caractérisé par la séquence informationnelle des acides nucléiques  
35 naturels, susceptible de s'hybrider à tout autre fragment nucléotidique dans des conditions prédéterminées,

l'enchaînement pouvant contenir des monomères de structures chimiques différentes et être obtenu à partir d'une molécule d'acide nucléique naturelle et/ou par recombinaison génétique et/ou par synthèse chimique ; un  
5 fragment nucléotidique peut être identique à un fragment génomique du virus MSRV-1 considéré par la présente invention, notamment un gène de ce dernier, par exemple pol ou env dans le cas dudit virus ;

- ainsi un monomère peut être un nucléotide  
10 naturel d'acide nucléique, dont les éléments constitutifs sont un sucre, un groupement phosphate et une base azotée; dans l'ARN le sucre est le ribose, dans l'ADN le sucre est le désoxy-2-ribose; selon qu'il s'agit de l'ADN ou l'ARN, la base azotée est choisie parmi l'adénine, la guanine,  
15 l'uracile, la cytosine, la thymine; ou le nucléotide peut être modifié dans l'un au moins des trois éléments constitutifs ; à titre d'exemple, la modification peut intervenir au niveau des bases, générant des bases modifiées telles que l'inosine, la méthyl-5-  
20 désoxycytidine, la désoxyuridine, la diméthylamino-5-désoxyuridine, la diamino-2,6-purine, la bromo-5-désoxyuridine et toute autre base modifiée favorisant l'hybridation; au niveau du sucre, la modification peut consister dans le remplacement d'au  
25 moins un désoxyribose par un polyamide (8), et au niveau du groupement phosphate, la modification peut consister dans son remplacement par des esters, notamment choisis parmi les esters de diphosphate, d'alkyl et arylphosphonate et de phosphorothioate,

30 - par "séquence informationnelle", on entend toute suite ordonnée de monomères, dont la nature chimique et l'ordre dans un sens de référence, constituent ou non une information fonctionnelle de même qualité que celle des acides nucléiques naturels,

35 - par hybridation, on entend le processus au cours duquel, dans des conditions opératoires appropriées,

deux fragments nucléotidiques, ayant des séquences suffisamment complémentaires, s'apparient pour former une structure complexe, notamment double ou triple, de préférence sous forme d'hélice,

- 5           - une sonde comprend un fragment nucléotidique synthétisé par voie chimique ou obtenu par digestion ou coupure enzymatique d'un fragment nucléotidique plus long, comprenant au moins six monomères, avantageusement de 10 à 100 monomères, de préférence 10 à 30 monomères, et
- 10 possédant une spécificité d'hybridation dans des conditions déterminées ; de préférence, une sonde possédant moins de 10 monomères n'est pas utilisée seule, mais l'est en présence d'autres sondes de taille aussi courte ou non ; dans certaines conditions particulières,
- 15 il peut être utile d'utiliser des sondes de taille supérieure à 100 monomères ; une sonde peut notamment être utilisée à des fins de diagnostic et il s'agira par exemple de sondes de capture et/ou de détection,

- la sonde de capture peut être immobilisée sur
- 20 un support solide par tout moyen approprié, c'est-à-dire directement ou indirectement, par exemple par covalence ou adsorption passive,

- la sonde de détection peut être marquée au moyen d'un marqueur choisi notamment parmi les isotopes
- 25 radioactifs, des enzymes notamment choisis parmi la peroxydase et la phosphatase alcaline et ceux susceptibles d'hydrolyser un substrat chromogène, fluorigène ou luminescent, des composés chimiques chromophores, des composés chromogènes, fluorigènes ou luminescents, des
- 30 analogues de bases nucléotidiques, et la biotine,

- les sondes utilisées à des fins de diagnostic de l'invention peuvent être mises en oeuvre dans toutes les techniques d'hybridation connues, et notamment les techniques dites "DOT-BLOT" (9), "SOUTHERN BLOT" (10),
- 35 "NORTHERN BLOT" qui est une technique identique à la technique "SOUTHERN BLOT" mais qui utilise de l'ARN comme

cible, la technique SANDWICH (11) ; avantageusement, on utilise la technique SANDWICH dans la présente invention, comprenant une sonde de capture spécifique et/ou une sonde de détection spécifique, étant entendu que la sonde de capture et la sonde de détection doivent présenter une séquence nucléotidique au moins partiellement différente,

- toute sonde selon la présente invention peut s'hybrider in vivo ou in vitro sur l'ARN et/ou sur l'ADN, pour bloquer les phénomènes de réplication, notamment traduction et/ou transcription, et/ou pour dégrader ledit ADN et/ou ARN,

- une amorce est une sonde comprenant au moins six monomères, et avantageusement de 10 à 30 monomères, possédant une spécificité d'hybridation dans des conditions déterminées, pour l'initiation d'une polymérisation enzymatique, par exemple dans une technique d'amplification telle que la PCR (Polymerase Chain Reaction), dans un procédé d'élongation, tel que le séquençage, dans une méthode de transcription inverse ou analogue,

- deux séquences nucléotidiques ou peptidiques sont dites équivalentes ou dérivées l'une par rapport à l'autre, ou par rapport à une séquence de référence, si fonctionnellement les biopolymères correspondants peuvent jouer sensiblement le même rôle, sans être identiques, vis-à-vis de l'application ou utilisation considérée, ou dans la technique dans laquelle elles interviennent ; sont notamment équivalentes deux séquences obtenues du fait de la variabilité naturelle, notamment mutation spontanée de l'espèce à partir de laquelle elles ont été identifiées, ou induite, ainsi que deux séquences homologues, l'homologie étant définie ci-après,

- par "variabilité", on entend toute modification, spontanée ou induite d'une séquence, notamment par substitution, et/ou insertion, et/ou délétion de nucléotides et/ou de fragments nucléotidiques,

et/ou extension et/ou raccourcissement de la séquence à l'une au moins des extrémités; une variabilité non naturelle peut résulter des techniques de génie génétique utilisées, par exemple du choix des amorces de synthèse  
5 dégénérées ou non, retenues pour amplifier un acide nucléique; cette variabilité peut se traduire par des modifications de toute séquence de départ, considérée comme référence, et pouvant être exprimées par un degré d'homologie par rapport à ladite séquence de référence,

10 - l'homologie caractérise le degré d'identité de deux fragments nucléotidiques ou peptidiques comparés ; elle se mesure par le pourcentage d'identité qui est notamment déterminé par comparaison directe de séquences nucléotidiques ou peptidiques, par rapport à des séquences  
15 nucléotidiques ou peptidiques de référence,

- ce pourcentage d'identité a été spécifiquement déterminé pour les fragments nucléotidiques, notamment clones relevant de la présente invention, homologues aux fragments identifiés pour le virus MSRV-1 par SEQ ID N° 1  
20 à N° 9, SEQ ID NO 46, SEQ ID NO 51 à SEQ ID NO 53, SEQ ID NO 40, SEQ ID NO 56 et SEQ ID NO 57, ainsi que pour les sondes et amorces homologues aux sondes et amorces identifiées par SEQ ID NO 20 à SEQ ID NO 24, SEQ ID NO 26, SEQ ID NO 16 à SEQ ID NO 19, SEQ ID NO 31 à SEQ ID NO 33,  
25 SEQ ID NO 45, SEQ ID NO 47, SEQ ID NO 48, SEQ ID NO 49, SEQ ID NO 50, SEQ ID NO 55, SEQ ID NO 40, SEQ ID NO 56 et SEQ ID NO 57 ; à titre d'exemple, le plus faible pourcentage d'identité observé entre les différents consensus généraux en acides nucléiques obtenus à partir  
30 de fragments d'ARN viral de MSRV-1, issu des lignées LM7PC et PLI-2 selon un protocole détaillé plus loin, est de 67% dans la région décrite à la figure 1,

- tout fragment nucléotidique est dit équivalent ou dérivé d'un fragment de référence, s'il présente une  
35 séquence nucléotidique équivalente à la séquence du fragment de référence ; selon la définition précédente,



sont notamment équivalents à un fragment nucléotidique de référence :

(a) tout fragment susceptible de s'hybrider au moins partiellement avec le complément du fragment de  
5 référence,

(b) tout fragment dont l'alignement avec le fragment de référence conduit à mettre en évidence des bases contigues identiques, en nombre plus important qu'avec tout autre fragment provenant d'un autre groupe  
10 taxonomique,

(c) tout fragment résultant ou pouvant résulter de la variabilité naturelle de l'espèce, à partir de laquelle il est obtenu,

(d) tout fragment pouvant résulter des techniques  
15 de génie génétique appliquées au fragment de référence,

(e) tout fragment, comportant au moins huit nucléotides contigus, codant pour un peptide homologue ou identique au peptide codé par le fragment de référence,

(f) tout fragment différent du fragment de  
20 référence, par insertion, délétion, substitution d'au moins un monomère, extension, ou raccourcissement à l'une au moins de ses extrémités ; par exemple, tout fragment correspondant au fragment de référence, flanqué à l'une au moins de ses extrémités par une séquence nucléotidique ne  
25 codant pas pour un polypeptide,

- par polypeptide, on entend notamment tout peptide d'au moins deux acides aminés, notamment oligopeptide, protéine, extrait, séparé, ou substantiellement isolé ou synthétisé, par l'intervention  
30 de la main de l'homme, notamment ceux obtenus par synthèse chimique, ou par expression dans un organisme recombinant,

- par polypeptide codé de manière partielle par un fragment nucléotidique, on entend un polypeptide présentant au moins trois acides aminés codés par au moins  
35 neuf monomères contigus compris dans ledit fragment nucléotidique,

- un acide aminé est dit analogue à un autre acide aminé, lorsque leurs caractéristiques physico-chimiques respectives, telles que polarité, hydrophobicité, et/ou basicité, et/ou acidité, et/ou neutralité, sont sensiblement les mêmes ; ainsi, une leucine est analogue à une isoleucine.

- tout polypeptide est dit équivalent ou dérivé d'un polypeptide de référence, si les polypeptides comparés ont sensiblement les mêmes propriétés, et notamment les mêmes propriétés antigéniques, immunologiques, enzymologiques et/ou de reconnaissance moléculaire ; est notamment équivalent à un polypeptide de référence :

(a) tout polypeptide possédant une séquence dont au moins un acide aminé a été substitué par un acide aminé analogue,

(b) tout polypeptide ayant une séquence peptidique équivalente, obtenue par variation naturelle ou induite dudit polypeptide de référence, et/ou du fragment nucléotidique codant pour ledit polypeptide,

(c) un mimotope dudit polypeptide de référence,

(d) tout polypeptide dans la séquence duquel un ou plusieurs acides aminés de la série L sont remplacés par un acide aminé de la série D, et vice versa,

(e) tout polypeptide dans la séquence duquel on a introduit une modification des chaînes latérales des acides aminés, telle que par exemple une acétylation des fonctions amines, une carboxylation des fonctions thiol, une estérification des fonctions carboxyliques,

(f) tout polypeptide dans la séquence duquel une ou des liaisons peptidiques ont été modifiées, comme par exemple les liaisons carba, rétro, inverso, rétro-inverso, réduites, et méthylène-oxy,

(g) tout polypeptide dont au moins un antigène est reconnu par anticorps dirigé contre un polypeptide de référence,

- le pourcentage d'identité caractérisant l'homologie de deux fragments peptidiques comparés est selon la présente invention d'au moins 50% et de préférence au moins 70 %.

5 Etant donné qu'un virus possédant une activité enzymatique transcriptase inverse peut être génétiquement caractérisé aussi bien sous forme d'ARN que d'ADN, il sera fait mention aussi bien de l'ADN que de l'ARN viral pour caractériser les séquences relatives à un virus possédant  
10 une telle activité transcriptase inverse, dit MSRV-1 selon la présente description.

Les expressions d'ordre utilisées dans la présente description et les revendications, telles que "première séquence nucléotidique" ne sont pas retenues  
15 pour exprimer un ordre particulier, mais pour définir plus clairement l'invention.

Par détection d'une substance ou agent, on entend ci-après aussi bien une identification, qu'une quantification, ou une séparation ou isolement de ladite  
20 substance ou dudit agent.

L'invention sera mieux comprise à la lecture de la description détaillée qui va suivre faite en référence aux figures annexées dans lesquelles :

- la figure 1 représente des consensus généraux  
25 en acides nucléiques des clones MSRV-1B amplifiés par la technique PCR dans la région "pol" définie par Shih (12), à partir d'ADN viral issu des lignées LM7PC et PLI-2, et identifiés sous les références SEQ ID NO 3, SEQ ID NO 4, SEQ ID NO 5, et SEQ ID NO 6, et le consensus commun avec  
30 amorces d'amplification portant la référence SEQ ID NO 7,

- la figure 2 donne la définition d'une trame de lecture fonctionnelle pour chaque famille de type MSRV-1B/"PCR pol", lesdites familles A à D étant définies respectivement par les séquences nucléotidiques SEQ ID NO  
35 3, SEQ ID NO 4, SEQ ID NO 5, et SEQ ID NO 6, décrites à la figure 1,

- la figure 3 donne un exemple de consensus des séquences de MSRV-2B, identifié par SEQ ID NO 11,

- la figure 4 est une représentation de l'activité transcriptase inverse (RT) en dpm (désintégration par minute), dans les fractions de saccharose prélevées sur un gradient de purification des virions produits par les lymphocytes B en culture d'un patient atteint de SEP,

- la figure 5 donne, dans les mêmes conditions expérimentales qu'à la figure 4, le dosage de l'activité transcriptase inverse dans la culture d'une lignée de lymphocytes B obtenue à partir d'un témoin exempt de SEP,

- la figure 6 représente la séquence nucléotidique du clone PSJ17 (SEQ ID NO 9),

- la figure 7 représente la séquence nucléotidique SEQ ID NO 8, du clone dénommé M003-P004,

- la figure 8 représente la séquence nucléotidique SEQ ID NO 2 du clone F11-1 ; la partie repérée entre les deux flèches dans la région de l'amorce correspond à une variabilité imposée par le choix de l'amorce ayant servi au clonage de F11-1 ; sur cette même figure, la traduction en acides aminés est représentée,

- la figure 9 représente la séquence nucléotidique SEQ ID NO 1, et une trame fonctionnelle de lecture possible en acides aminés de SEQ ID NO 1 ; sur cette séquence, les séquences consensus du gène pol sont soulignées,

- les figures 10 et 11 donnent les résultats d'une PCR, sous forme de photographie sous lumière ultra-violette d'un gel d'agarose imprégné de bromure d'éthidium, des produits d'amplification obtenus à partir des amorces identifiées par SEQ ID NO 16, SEQ ID NO 17, SEQ ID NO 18, et SEQ ID NO 19.

- la figure 12 donne une représentation matricielle de l'homologie entre SEQ ID NO 1 de MSRV-1 et celle d'un rétrovirus endogène dénommé HSERV9 ; cette

homologie d'au moins 65 % est mise en évidence par un trait plein, l'absence de trait signifiant une homologie inférieure à 65 % ;

5 - la figure 13 représente la séquence nucléotidique SEQ ID NO 46 du clone FBd3,

- la figure 14 représente l'homologie de séquence entre le clone FBd3 et le rétrovirus HSERV-9 ;

- la figure 15 représente la séquence nucléotidique SEQ ID NO 51 du clone t pol ;

10 - les figures 16 et 17 représentent respectivement les séquences nucléotidiques SEQ ID NO 52 et SEQ ID NO 53, des clones JLBc1 et JLBc2 respectivement ;

15 - la figure 18 représente l'homologie de séquence entre le clone JLBc1 et le clone FBd3,

- et la figure 19 l'homologie de séquence entre le clone JLBc2 et le clone FBd3 ;

- la figure 20 représente l'homologie de séquence entre les clones JLBc1 et JLBc2 ;

20 - les figures 21 et 22 représentent l'homologie de séquence entre le rétrovirus HSERV-9, et respectivement les clones JLBc1 et JLBc2 ;

- la figure 23 représente la séquence nucléotidique SEQ ID NO 56 du clone GM3 ;

25 - la figure 24 représente l'homologie de séquence entre le rétrovirus HSERV-9 et le clone GM3 ;

- la figure 25 représente la localisation des différents clones étudiés, par rapport au génome du rétrovirus connu ERV9 ;

30 - la figure 26 représente la position des clones F11-1, M003-P004, MSRV-1B et PSJ17 dans la région dénommée ci-après MSRV-1 pol \* ;

- la figure 27, éclatée en trois figures successives 27a, 27b, 27c représente un cadre de lecture  
35 possible, couvrant l'ensemble du gène pol ;

- la figure 28 représente selon SEQ ID NO 40, la séquence nucléotidique codant pour le fragment peptidique POL2B, ayant la séquence en acides aminés identifiée par SEQ ID NO 39 ;

5           - la figure 29 représente les valeurs de DO (tests ELISA) à 492 nm obtenues pour 29 sérums de patients SEP et 32 sérums de témoins sains testés avec un anticorps anti-IgG ;

10           - la figure 30 représente les valeurs de DO (tests ELISA) à 492 nm obtenues pour 36 sérums de patients SEP et 42 sérums de témoins sains testés avec un anticorps anti-IgM ;

15           - les figures 31 à 33 représentent les résultats obtenus (intensité relative des spots) pour 43 octapeptides chevauchants couvrant la séquence en acides aminés 61-110, selon la technique SpotsScan, respectivement avec un pool de sérums de SEP, avec un pool de sérums témoins et avec le pool de sérums SEP après déduction d'un bruit de fond correspondant au signal maximum détecté sur au moins un octapeptide avec le sérum témoin (intensité = 1), étant entendu que ces sérums ont été dilués au 1/50ème. La barre à l'extrême droite représente un standard d'échelle graphique sans rapport avec le test sérologique ;

25           - la figure 34 représente les SEQ ID NO 41 et SEQ ID NO 42 de deux polypeptides comprenant immuno-dominantes, tandis que SEQ ID NO 43 et 44 représentent des polypeptides immuno-réactifs et spécifiques à la SEP ;

30           - la figure 35 représente la séquence nucléotidique SEQ ID NO 59 du clone LB19 et trois trames de lecture potentielles en acides aminés de SEQ ID NO 59 ;

35           - la figure 36 représente la séquence nucléotidique SEQ ID NO 88 (GAG\*) et une trame de lecture potentielle en acides aminés de SEQ ID NO 88 ;

- la figure 37 représente l'homologie de séquence entre le clone FBd13 et le rétrovirus HSERV-9 ; selon cette représentation, le trait plein signifie un pourcentage d'homologie supérieur ou égal à 70% et  
5 l'absence de trait signifie un pourcentage d'homologie inférieur ;

- la figure 38 représente la séquence nucléotidique SEQ ID NO 61 du clone FP6 et trois trames de lecture potentielles en acides aminés de SEQ ID NO 61 ;

10 - la figure 39 représente la séquence nucléotidique SEQ ID NO 89 du clone G+E+A et trois trames de lecture potentielles en acides aminés de SEQ ID NO 89 ;

- la figure 40 représente un cadre de lecture trouvé dans la région E et codant pour une protéase  
15 rétrovirale MSRV-1 identifiée par SEQ ID NO 90 ;

- la figure 41 représente la réponse de chaque sérum de patients atteints de SEP indiqué par le symbole (+) et de patients sains symbolisé par (-), testé avec un anticorps anti-IgG exprimée en densité optique nette à 492  
20 nm;

- la figure 42 représente la réponse de chaque sérum de patients atteints de SEP indiqué par les symboles (+) et (QS) et de patients sains (-), testé avec un anticorps anti-IgM exprimée en densité optique nette à 492  
25 nm.

**EXEMPLE 1 : OBTENTION DE CLONES DENOMMES MSRV-1B ET MSRV-2B, DEFINISSANT RESPECTIVEMENT UN RETROVIRUS MSRV-1 ET UN AGENT CO-INFECTANT MSRV2, PAR AMPLIFICATION  
30 PCR "NICHEE" DES REGIONS POL CONSERVEES DES RETROVIRUS, SUR DES PREPARATIONS DE VIRIONS ISSUES DES LIGNEES LM7PC ET PLI-2**

Une technique PCR dérivée de la technique publiée  
35 par Shih (12) a été utilisée. Cette technique permet, par traitement de tous les composants du milieu réactionnel

par la DNase, d'éliminer toute trace d'ADN contaminant. Elle permet parallèlement, par l'utilisation d'amorces différentes mais chevauchantes dans deux séries successives de cycles d'amplifications PCR, d'augmenter  
5 les chances d'amplifier un ADNc synthétisé à partir d'une quantité d'ARN faible au départ et encore réduite dans l'échantillon par l'action parasite de la DNase sur l'ARN. En effet, la DNase est utilisée dans des conditions d'activité en excès qui permettent d'éliminer toute trace  
10 d'ADN contaminant, avant inactivation de cette enzyme restant dans l'échantillon par chauffage à 85°C pendant 10 minutes. Cette variante de la technique PCR décrite par Shih (12), a été utilisée sur un ADNc synthétisé à partir des acides nucléiques de fractions de particules  
15 infectantes purifiées sur gradient de saccharose, selon la technique décrite par H. Perron (13), à partir de l'isolat "POL-2" (ECACC n°V92072202) produit par la lignée PLI-2 (ECACC n°92072201) d'une part, et à partir de l'isolat MS7PG (ECACC n°V93010816) produit par la lignée LM7PC  
20 (ECACC n°93010817) d'autre part. Ces cultures ont été obtenues selon les méthodes ayant fait l'objet des demandes de brevet publiées sous les n° WO 93/20188 et WO 93/20189.

Après clonage avec le TA Cloning Kit® des  
25 produits amplifiés par cette technique et analyse de la séquence à l'aide d'un séquenceur automatique "Automatic Sequencer, modèle 373A" d'Applied Biosystems, les séquences ont été analysées à l'aide du logiciel Geneworks®, sur la dernière version disponible de la  
30 banque de données Genbank®.

Les séquences clonées et séquencées à partir de ces échantillons correspondent notamment à deux types de séquences: un premier type de séquence, retrouvé dans la majorité des clones (55 % des clones issus des isolats  
35 POL-2 des culture PLI-2, et, 67 % des clones issus des isolats MS7PG des cultures LM7PC), qui correspond à une



famille de séquences "pol" proches mais différentes du rétrovirus endogène humain dénommé ERV-9 ou HSERV-9, et un second type de séquence qui correspond à des séquences très fortement homologues à une séquence attribuée à un  
5 autre agent infectant et/ou pathogène dénommé MSRV-2.

Le premier type de séquences représentant la majorité des clones est constituée de séquences dont la variabilité permet de définir quatre sous-familles de séquences. Ces sous-familles sont suffisamment proches  
10 entre elles pour qu'on puisse les considérer comme des quasi-espèces provenant d'un même rétrovirus, tel que cela est bien connu pour le rétrovirus HIV-1 (14), ou comme le résultat de l'interférence avec plusieurs provirus endogènes co-régulés dans les cellules productrices. Ces  
15 éléments endogènes plus ou moins défectifs sont sensibles aux mêmes signaux de régulation générés éventuellement par un provirus réplicatif, puisqu'ils appartiennent à la même famille de rétrovirus endogènes (15). Cette nouvelle  
20 famille de rétrovirus endogènes ou, alternativement, cette nouvelle espèce rétrovirale dont on a obtenu en culture la génération de quasi-espèces, et qui contient un consensus des séquences décrites ci-dessous est dénommée MSRV-1B.

Dans la figure 1 sont présentés les consensus généraux des séquences des différents clones MSRV-1B  
25 séquencés lors de cette expérience, ces séquences étant respectivement identifiées par SEQ ID NO 3, SEQ ID NO 4, SEQ ID NO 5, et SEQ ID NO 6. Ces séquences présentent une homologie en acides nucléiques allant de 70 % à 88 % avec la séquence HSERV9 référencée X57147 et M37638 dans la  
30 base de données Genbank®. Quatre séquences nucléiques "consensus" représentatives de différentes quasi-espèces d'un rétrovirus MSRV-1B éventuellement exogène, ou de différentes sous-familles d'un rétrovirus endogène MSRV-1B, ont été définies. Ces consensus représentatifs  
35 sont présentés dans la figure 2, avec la traduction en acides aminés. Une trame de lecture fonctionnelle existe

pour chaque sous-famille de ces séquences MSRV-1B, et l'on peut voir que la trame de lecture ouverte fonctionnelle correspond à chaque fois à la séquence en acides aminés venant en deuxième ligne sous la séquence en acide  
5 nucléiques. Le consensus général de la séquence MSRV-1B, identifié par SEQ ID NO 7 et obtenu par cette technique PCR dans la région "pol" est présenté dans la figure 1.

Le deuxième type de séquence représentant la majorité des clones séquencés est représenté par la  
10 séquence MSRV-2B présentée dans la figure 3 et identifiée par SEQ ID NO 11. Les différences observées dans les séquences correspondant aux amorces PCR s'expliquent par l'utilisation d'amorces dégénérées en mélange utilisées dans des conditions techniques différentes.

15 La séquence MSRV-2B (SEQ ID NO 11) est suffisamment divergente des séquences rétrovirales déjà décrites dans les banques de données pour qu'on puisse avancer qu'il s'agit d'une région de séquence appartenant à un nouvel agent infectant, dénommé MSRV-2. Cet agent  
20 infectant s'apparenterait à priori, d'après l'analyse des premières séquences obtenues, à un rétrovirus, mais, étant donnée la technique utilisée pour obtenir cette séquence, il pourrait aussi s'agir d'un virus à ADN dont le génome code pour une enzyme qui possède accessoirement une  
25 activité transcriptase inverse comme c'est le cas, par exemple, pour le virus de l'hépatite B, HBV (12). De plus, le caractère aléatoire des amorces dégénérées utilisées pour cette technique d'amplification PCR peut très bien avoir permis, du fait d'homologies de séquences imprévues  
30 ou de sites conservés dans le gène d'une enzyme apparentée, l'amplification d'un acide nucléique provenant d'un agent pathogène et/ou co-infectant prokaryote ou eukaryote (protiste).

**EXEMPLE 2 : OBTENTION DE CLONES DENOMMES MSRV-1B ET MSRV-2B, DEFINISSANT UNE FAMILLE MSRV-1 et MSRV2, PAR AMPLIFICATION PCR "NICHEE" DES REGIONS POL CONSERVEES DES RETROVIRUS, SUR DES PREPARATIONS DE LYMPHOCYTES B D'UN**  
5 **NOUVEAU CAS DE SEP**

La même technique PCR modifiée d'après la technique de Shih (12) a été utilisée pour amplifier et séquencer le matériel nucléaire ARN présent dans une  
10 fraction de virions purifiée au pic d'activité transcriptase inverse "de type LM7", sur gradient de saccharose selon la technique décrite par H. Perron (13), et selon les protocoles mentionnés dans l'exemple 1, à partir d'une lignée lymphoblastoïde spontanée, obtenue par  
15 auto-immortalisation en culture de lymphocytes B d'un patient SEP séropositif pour le virus d'Epstein-Barr (EBV), après mise en culture des cellules lymphoïdes sanguines dans un milieu de culture approprié contenant une concentration appropriée de cyclosporine A.  
20 Une représentation de l'activité transcriptase inverse dans les fractions de saccharose prélevées sur un gradient de purification des virions produits par cette lignée est présentée dans la figure 4. De même, les surnageants de culture d'une lignée B obtenue dans les mêmes conditions à  
25 partir d'un témoin exempt de SEP ont été traités dans les mêmes conditions, et le dosage de l'activité transcriptase inverse dans les fractions du gradient de saccharose s'est avéré négatif partout (bruit de fond) et est présenté dans la figure 5. La fraction 3 du gradient correspondant à la  
30 lignée B de SEP et la même fraction sans activité transcriptase inverse du gradient témoin non-SEP, ont été analysées par la même technique RT-PCR que précédemment, dérivée de Shih (12), suivie des mêmes étapes de clonage et de séquençage tels que décrites dans l'exemple 1.  
35 Il est tout à fait notable que les séquences de type MSRV-1 et MSRV-2 soient retrouvées dans le seul

matériel associé à un pic d'activité transcriptase inverse "de type LM7" provenant de la lignée lymphoblastoïde B de SEP. Ces séquences n'ont pas été retrouvées avec le matériel de la lignée lymphoblastoïde B témoin (non-SEP) dans 26 clones recombinants pris au hasard. Seules les séquences contaminantes de type Mo-MuLV provenant de la transcriptase inverse commerciale utilisée pour l'étape de synthèse de l'ADNc et des séquences sans analogie rétrovirale particulière ont été retrouvées chez ce témoin, du fait de l'amplification "consensus" des séquences de polymérases homologues que réalise cette technique PCR. De plus, l'absence de cible concentrée qui fait compétition pour la réaction d'amplification dans l'échantillon témoin permet l'amplification de contaminants dilués. La différence de résultats est à l'évidence hautement significative ( $\chi^2$ ,  $p < 0,001$ ).

**EXEMPLE 3 : OBTENTION D'UN CLONE PSJ17, DEFINISSANT UN RETROVIRUS MSRV-1, PAR REACTION DE TRANSCRIPTASE INVERSE ENDOGENE SUR UNE PREPARATION DE VIRION ISSUE DE LA LIGNEE PLI-2.**

Cette approche vise à obtenir des séquences d'ADN rétrotranscrites à partir de l'ARN supposé rétroviral dans l'isolat, en utilisant l'activité transcriptase inverse présente dans ce même isolat. Cette activité transcriptase inverse ne peut théoriquement fonctionner qu'en présence d'un ARN rétroviral, lié à un ARNt amorce ou hybridé à des brins courts d'ADN déjà rétro-transcrits dans les particules rétrovirales (16). Ainsi l'obtention de séquences rétrovirales spécifiques dans un matériel contaminé par des acides nucléiques cellulaires, a été optimisée selon ces auteurs grâce à l'amplification enzymatique spécifique des portions d'ARN viraux par une activité transcriptase inverse virale. Les auteurs ont

pour cela, déterminé les conditions physico-chimiques particulières dans lesquelles cette activité enzymatique de transcription inverse sur des ARN contenus dans des virions pouvait être effective in vitro. Ces conditions  
5 correspondent à la description technique des protocoles présentés ci-dessous (réaction de RT endogène, purification, clonage et séquençage).

L'approche moléculaire a consisté à utiliser une préparation de virion concentré, mais non purifié, obtenu  
10 à partir des surnageants de culture de la lignée PLI-2 préparés selon la méthode suivante : les surnageants de culture sont collectés deux fois par semaine, pré-centrifugés à 10 000 trs/min pendant 30 minutes pour éliminer les débris cellulaires, et ensuite, congelés  
15 à -80°C ou utilisés tels quels pour les étapes suivantes. Les surnageants frais ou décongelés sont centrifugés sur un coussin de PBS-glycérol 30% à 100 000g (ou 30 000 trs/min dans un rotor LKB-HITACHI de type 45 T) pendant 2h à 4°C. Après élimination du surnageant, le  
20 culot sédimenté est repris dans un petit volume de PBS et constitue la fraction de virion concentré, mais non purifié. Cet échantillon viral concentré mais non purifié a été utilisé pour effectuer une réaction dite de transcription inverse endogène, telle que décrite  
25 ci-après.

Un volume de 200 µl de virion purifié selon le protocole décrit ci-dessus et contenant une activité transcriptase inverse d'environ 1-5 millions de dpm est décongelé à 37°C jusqu'à apparition d'une phase liquide,  
30 puis placé sur de la glace. Un tampon 5 fois concentré a été préparé avec les composants suivants: Tris-HCl pH 8,2, 500 mM; NaCl 75 mM; MgCl<sub>2</sub> 25 mM; DTT 75 mM et NP 40 0,10 % ; 100 µl de tampon 5X + 25 µl d'une solution de dATP 100mM + 25 µl d'une solution de dTTP 100 mM + 25 µl  
35 d'une solution de dGTP 100 mM + 25 µl d'une solution de dCTP 100 mM + 100 µl d'eau distillée stérile + 200 µl de

la suspension de virions (activité T.I. de 5 millions de DPM) dans le PBS ont été mélangés et incubés à 42°C pendant 3 heures. Après cette incubation le mélange réactionnel est directement ajouté à un mélange tamponné  
5 phénol/chloroforme/alcool isoamylique (Sigma ref. P 3803); la phase aqueuse est collectée et un volume d'eau distillée stérile est ajouté à la phase organique pour ré-extraire le matériel nucléaire résiduel. Les phases aqueuses collectées sont regroupées et les acides  
10 nucléiques contenus sont précipités par addition d'acétate de sodium 3M, pH 5,2 au 1/10 de volume + 2 volumes d'éthanol + 1 µl de glycogène (Boehringer-Mannheim ref. 901 393) et mise à -20°C de l'échantillon pendant 4h ou la nuit à +4°C. Le précipité obtenu après  
15 centrifugation est ensuite lavé avec de l'éthanol à 70% et resuspendu dans 60 ml d'eau distillée. Les produits de cette réaction ont ensuite été purifiés, clonés et séquencés selon le protocole décrit ci-après: des ADN bouts francs avec des adénines non-appariées aux  
20 extrémités ont été générés: une réaction de "remplissage" a d'abord été effectuée: 25 µl de la solution d'ADN précédemment purifiée ont été mélangés avec 2 µl d'une solution 2,5 mM contenant, en quantité équimolaire, dATP + dGTP + dTTP + dCTP / 1 µl d'ADN polymérase T4  
25 (Boehringer-Mannheim ref. 1004 786) / 5 µl de 10X "incubation buffer for restriction enzyme" (Boehringer-Mannheim ref. 1417 975) / 1 µl d'une solution à 1 % de sérum-albumine bovine / 16 µl d'eau distillée stérile. Ce mélange a été incubé 20 minutes à 11°C. 50 µl de tampon TE  
30 et 1 µl de glycogène (Boehringer-Mannheim ref. 901 393) y ont été ajoutés avant extraction des acides nucléiques avec du phénol/chloroforme/alcool isoamylique (Sigma ref. P 3803), et précipitation à l'acétate de sodium comme décrit précédemment. L'ADN précipité après  
35 centrifugation est resuspendu dans 10 µl de tampon 10 mM Tris pH 7,5. Puis, 5 µl de cette suspension ont été

mélangés avec 20  $\mu$ l de tampon Taq 5X, 20  $\mu$ l de 5mM dATP, 1  $\mu$ l (5U) de Taq ADN polymérase (Amplitaq<sup>TM</sup>) et 54  $\mu$ l d'eau distillée stérile. Ce mélange est incubé 2 h à 75°C avec un film d'huile à la surface de la solution. L'ADN en suspension dans la solution aqueuse prélevée en dessous du film d'huile après incubation est précipité comme décrit précédemment et resuspendu dans 2  $\mu$ l d'eau distillée stérile. L'ADN obtenu a été inséré dans un plasmide à l'aide du kit TA Cloning<sup>TM</sup>. Les 2  $\mu$ l de solution d'ADN ont été mélangés avec 5  $\mu$ l d'eau distillée stérile, 1  $\mu$ l d'un tampon de ligation 10 fois concentré "10X LIGATION BUFFER", 2  $\mu$ l de "pCR<sup>TM</sup> VECTOR" (25 ng/ml) et 1  $\mu$ l de "TA DNA LIGASE". Ce mélange a été incubé une nuit à 12°C. Les étapes suivantes ont été réalisées conformément aux instructions du kit TA Cloning® (British Biotechnology). A la fin de la procédure, les colonies blanches de bactéries recombinantes (white) ont été repiquées pour être cultivées et permettre l'extraction des plasmides incorporés selon la procédure dite de "miniprep" (17). La préparation de plasmide de chaque colonie recombinante a été coupée par une enzyme de restriction appropriée et analysée sur gel d'agarose. Les plasmides possédant un insert détecté sous lumière UV après marquage du gel au bromure d'éthidium, ont été sélectionnés pour le séquençage de l'insert, après hybridation avec une amorce complémentaire du promoteur Sp6 présent sur le plasmide de clonage du TA cloning kit®. La réaction préalable au séquençage a ensuite été effectuée selon la méthode préconisée pour l'utilisation du kit de séquençage "Prism ready reaction kit dye deoxyterminator cycle sequencing kit" (Applied Biosystems, ref. 401384) et le séquençage automatique a été réalisé avec l'appareil "Automatic Sequencer, modèle 373 A" Applied Biosystems, selon les instructions du fabricant.

L'analyse discriminante sur les banques de données informatiques des séquences clonées à partir des

fragments d'ADN présents dans le mélange réactionnel a permis de mettre en évidence une séquence de type rétroviral. Le clone correspondant PSJ17 a été entièrement séquencé et la séquence obtenue, présentée dans la figure 6 et identifiée apr SEQ ID N°9, a été analysée à l'aide du logiciel "Geneworks®" sur les banques de données actualisées "Genebank®". L'analyse des banques de données n'a pas permis de trouver de séquence identique déjà décrite. Seule une homologie partielle a pu être retrouvée avec certains éléments rétroviraux connus. L'homologie relative la plus intéressante concerne un rétrovirus endogène dénommé ERV-9, ou HSERV-9, selon les références (18).

15

**EXEMPLE 4 : AMPLIFICATION PCR DE LA SEQUENCE NUCLEIQUE CONTENUE ENTRE LA REGION 5' DEFINIE PAR LE CLONE "POL MSRV-1B" ET LA REGION 3' DEFINIE PAR LE CLONE PSJ17.**

20

Cinq oligonucléotides, M001, M002-A, M003-BCD, P004 et P005, ont été définis pour amplifier de l'ARN provenant de virions purifiés POL-2. Des réactions de contrôle ont été effectuées de manière à contrôler la présence de contaminants (réaction sur de l'eau).

25

L'amplification consiste en une étape de RT-PCR selon le protocole décrit dans l'Exemple 2, suivie d'une PCR "nichée" selon le protocole PCR décrit dans le document EP-A-0 569 272. Dans le premier cycle RT-PCR, les amorces M001 et P004 ou P005 sont utilisées. Dans le deuxième cycle PCR, les amorces M002-A ou M003-BCD, et l'amorce P004 sont utilisées. Les amorces sont positionnées comme suit :

35

	M002-A			
	M003-BCD			
M001	—	P004	P005	



\_\_\_\_\_ ARN  
 POL-2

<-----> <----->

5 pol MSRV-1B PSJ17

Leur composition est :

amorce M001: GGTCITICCAIGG (SEQ ID NO 20)

amorce M002-A: TTAGGGATAGCCCTCATCTCT (SEQ ID NO 21)

amorce M003-BCD: TCAGGGATAGCCCCCATCTAT (SEQ ID NO 22)

10 amorce P004: AACCCCTTGCCACTACATCAATTT (SEQ ID NO 23)

amorce P005: GCGTAAGGACTCCTAGAGCTATT (SEQ ID NO 24)

Le produit d'amplification "nichée" obtenu et dénommé M003-P004 est présenté dans la figure 7, et correspond à la séquence SEQ ID NO 8.

15

**EXEMPLE 5 : AMPLIFICATION ET CLONAGE D'UNE PORTION DU GENOME RETROVIRAL MSRV-1 A L'AIDE D'UNE SEQUENCE DEJA IDENTIFIEE, DANS UN ECHANTILLON DE VIRUS PURIFIE AU PIC D'ACTIVITE TRANSCRIPTASE INVERSE**

20

Une technique PCR dérivée de la technique publiée par Frohman (19) a été utilisée. La technique dérivée permet, à l'aide d'une amorce spécifique en 3' du génome à amplifier, d'élonguer la séquence vers la région 5' du génome à analyser. Cette variante technique est décrite dans la documentation de la firme "Clontech Laboratories Inc., (Palo-Alto California, USA) fournie avec son produit "5'-AmpliFINDER™ RACE Kit", qui a été utilisé sur une fraction de virion purifié comme décrit précédemment.

25

30

Les amorces 3' spécifiques utilisées dans le protocole du kit pour la synthèse du cDNA et l'amplification PCR sont respectivement complémentaires aux séquences MSRV-1 suivantes :

35 cDNA : TCATCCATGTACCGAAGG (SEQ ID N°25)  
 amplification : ATGGGGTTCCCAAGTTCCT (SEQ ID N°26)

Les produits issus de la PCR ont été purifiés après purification sur gel d'agarose selon les méthodes conventionnelles (17), puis resuspendus dans 10 ml d'eau distillée. Une des propriétés de la polymérase Taq consistant à ajouter une adénine à l'extrémité 3' de chacun des deux brins d'ADN, l'ADN obtenu a été directement inséré dans un plasmide à l'aide du kit TA Cloning™ (British Biotechnology). Les 2 µl de solution d'ADN ont été mélangés avec 5 µl d'eau distillée stérile, 1 µl d'un tampon de ligation 10 fois concentré "10X LIGATION BUFFER", 2 µl de "PCR™ VECTOR" (25 ng/ml) et 1 µl de "TA DNA LIGASE". Ce mélange a été incubé une nuit à 12°C. Les étapes suivantes ont été réalisées conformément aux instructions du kit TA Cloning® (British Biotechnology). A la fin de la procédure, les colonies blanches de bactéries recombinantes (white) ont été repiquées pour être cultivées et permettre l'extraction des plasmides incorporés selon la procédure dite de "miniprep" (17). La préparation de plasmide de chaque colonie recombinante a été coupée par une enzyme de restriction appropriée et analysée sur gel d'agarose. Les plasmides possédant un insert détecté sous lumière UV après marquage du gel au bromure d'éthidium, ont été sélectionnés pour le séquençage de l'insert, après hybridation avec une amorce complémentaire du promoteur Sp6 présent sur le plasmide de clonage du TA Cloning Kit®. La réaction préalable au séquençage a ensuite été effectuée selon la méthode préconisée pour l'utilisation du kit de séquençage "Prism ready reaction kit dye deoxyterminator cycle sequencing kit" (Applied Biosystems, ref. 401384) et le séquençage automatique a été réalisé avec l'appareil "Automatic Sequencer, modèle 373 A" Applied Biosystems selon les instructions du fabricant.

Cette technique a d'abord été appliquée à deux fractions de virion purifié comme décrit ci-après, sur

saccharose à partir de l'isolat "POL-2" produit par la lignée PLI-2 d'une part, et à partir de l'isolat MS7PG produit par la lignée LM7PC d'autre part. Les surnageants de cultures sont collectés deux fois par semaine, 5 pré-centrifugés à 10 000 trs/min pendant 30 minutes pour éliminer les débris cellulaires, et ensuite, congelés à -80°C ou utilisés tels quels pour les étapes suivantes. Les surnageants frais ou décongelés sont centrifugés sur un coussin de PBS-glycérol 30% à 100 000g (ou 10 30 000 trs/min dans un rotor LKB-HITACHI de type 45 T) pendant 2 h à 4°C. Après élimination du surnageant, le culot sédimenté est repris dans un petit volume de PBS et constitue la fraction de virions concentrés, mais non purifiés. Le virus concentré est ensuite déposé sur un 15 gradient de sucrose dans un tampon PBS stérile (15 à 50 % poids/poids) et ultracentrifugé à 35 000 trs/min (100 000g) pendant 12 h à +4°C dans un rotor à godets. 10 fractions sont recueillies et 20 µl sont prélevés dans chaque fraction après homogénéisation pour y doser 20 l'activité transcriptase inverse selon la technique décrite par H. Perron (3). Les fractions contenant le pic d'activité RT "de type LM7" sont ensuite diluées dans du tampon PBS stérile et ultracentrifugées une heure à 35 000 trs/min (100 000g) pour sédimenter les particules 25 virales. Le culot de virion purifié ainsi obtenu est alors repris dans un petit volume d'un tampon adéquat pour l'extraction d'ARN. La réaction de synthèse de cDNA mentionnée plus haut est réalisée sur cet ARN extrait de 30 la technique citée plus-haut a permis d'obtenir le clone F1-11 dont la séquence identifiée par SEQ ID NO 2, est présentée dans la figure 8.

Ce clone permet de définir, avec les différents clones préalablement séquencés, une région de longueur 35 importante (1,2 kb) représentative du gène "pol" du rétrovirus MSRV-1, telle que présentée dans la figure 9.

Cette séquence dénommée SEQ ID NO 1 est reconstituée à partir de différents clones se recouvrant à leurs extrémités, en corrigeant les artéfacts liés aux amorces et aux techniques d'amplification ou de clonage, qui interrompraient artificiellement la trame de lecture de l'ensemble. Cette séquence sera identifiée ci-après sous la dénomination "région MSRV-1 pol \*". Son degré d'homologie avec la séquence HSERV-9 est représenté à la figure 12.

10 Dans la figure 9, la trame de lecture potentielle avec sa traduction en acides aminés est présentée en-dessous de la séquence en acides nucléiques.

15 **EXEMPLE 6 : DETECTION DE SEQUENCES SPECIFIQUES MSRV-1 et MSRV-2 DANS DIFFERENTS ECHANTILLONS DE PLASMA PROVENANT DE PATIENTS ATTEINTS DE SEP OU DE TEMOINS.**

Une technique PCR a été utilisée pour détecter les génomes MSRV-1 et MSRV-2 dans des plasmas obtenus après prélèvement de sang sur EDTA de patients atteints de SEP et de témoins non-SEP.

20 L'extraction des ARN de plasma a été effectuée selon la technique décrite par P. Chomzynski (20), après addition d'un volume du tampon contenant du guanidinium thiocyanate à 1 ml de plasma gardé congelé à -80°C après recueil.

Pour MSRV-2, la PCR a été effectuée dans les mêmes conditions et avec les amorces suivantes :

- 30 - amorce 5', identifiée par SEQ ID NO 14  
5' GTAGTTCGATGTAGAAAGCG 3' ;  
- amorce 3', identifiée par SEQ ID NO 15  
5' GCATCCGGCAACTGCACG 3'.

Cependant des résultats similaires ont aussi été obtenus avec les amorces PCR suivantes dans deux

amplifications successives par PCR "nichée" sur échantillons d'acides nucléiques non-traités par la DNase.

Les amorces utilisées pour cette première étape de 40 cycles avec une température d'hybridation de 48°C sont les suivantes :

- amorce 5', identifiée par SEQ ID NO 27

5' GCCGATATCACCCGCCATGG 3', correspondant à une amorce PCR MSRV-2 5', pour une première PCR sur prélèvement de patients,

10 -amorce 3', identifiée par SEQ ID NO 28

5' GCATCCGGCAACTGCACG 3', correspondant à une amorce PCR MSRV-2 3', pour une première PCR sur prélèvement de patients.

Après cette étape, 10µl du produit d'amplification sont prélevés et utilisés pour réaliser une deuxième amplification PCR dite "nichée" avec des amorces situées à l'intérieur de la région déjà amplifiée. Cette deuxième étape se déroule sur 35 cycles, avec une température d'hybridation des amorces ("annealing") de 20 50°C. Le volume réactionnel est de 100µl.

Les amorces utilisées pour cette deuxième étape sont les suivantes :

- amorce 5', identifiée par SEQ ID NO 29

5' CGCGATGCTGGTTGGAGAGC 3', correspondant à une amorce PCR MSRV-2 5', pour une PCR-nichée sur prélèvement de patients,

- amorce 3', identifiée par SEQ ID NO 30

5' TCTCCACTCCGAATATTCCG 3', correspondant à une amorce PCR MSRV-2 3', pour une PCR-nichée sur prélèvement de patients.

Pour MSRV-1, l'amplification a été effectuée en deux étapes. De plus, l'échantillon d'acides nucléiques est préalablement traité par la DNase et un contrôle PCR sans RT (transcriptase inverse AMV) est effectué sur les 35 deux étapes d'amplification, de manière à vérifier que l'amplification RT-PCR provient exclusivement de l'ARN

MSRV-1. En cas de contrôle sans RT positif, l'échantillon aliquoté d'ARN de départ est à nouveau traité par la DNase et amplifié à nouveau.

Le protocole de traitement par la DNase dépourvue d'activité RNase est le suivant: l'ARN extrait est aliquoté en présence de "RNase inhibitor" (Boehringer-Mannheim) dans de l'eau traitée au DEPC à une concentration finale de 1 µg dans 10 µl; à ces 10 µl, est ajouté 1 µl de "RNase-free DNase" (Boehringer-Mannheim) et 1,2 µl de tampon à pH 5 contenant 0,1 M/l d'acétate de sodium et 5 mM/l de MgSO<sub>4</sub>; le mélange est incubé 15 min. à 20°C et porté à 95°C pendant 1,5 min. dans un "thermocycler".

La première étape de RT-PCR MSRV-1 est effectuée selon une variante du procédé d'amplification d'ARN tel que décrit dans la demande de brevet n° EP-A-0 569 272. Notamment, l'étape de synthèse d'ADNc est effectuée à 42°C pendant une heure; l'amplification PCR se déroule sur 40 cycles, avec une température d'hybridation des amorces ("annealing") de 53°C. Le volume réactionnel est de 100 µl.

Les amorces utilisées pour cette première étape sont les suivantes :

- amorce 5', identifiée par SEQ ID NO 16  
5' AGGAGTAAGGAAACCCAACGGAC 3' ;
- amorce 3', identifiée par SEQ ID NO 17  
5' TAAGAGTTGCACAAGTGCG 3'.

Après cette étape, 10 µl du produit d'amplification sont prélevés et utilisés pour réaliser une deuxième amplification PCR dite "nichée" avec des amorces situées à l'intérieur de la région déjà amplifiée. Cette deuxième étape se déroule sur 35 cycles, avec une température d'hybridation des amorces ("annealing") de 53°C. Le volume réactionnel est de 100 µl.

Les amorces utilisées pour cette deuxième étape sont les suivantes :

- amorce 5', identifiée par SEQ ID NO 18

5' TCAGGGATAGCCCCCATCTAT 3' ;  
- amorce 3', identifiée par SEQ ID NO 19  
5' AACCCCTTTGCCACTACATCAATTT 3'.

Dans les figures 10 et 11, sont présentés les  
5 résultats de PCR sous forme de photographies sous lumière  
ultra-violette de gels d'agarose imprégnés de bromure  
d'éthidium, dans lesquels on a effectué une électrophorèse  
des produits d'amplification PCR déposés séparément dans  
les différents puits.

10 La photographie du haut (figure 10) représente le  
résultat de l'amplification spécifique MSRV-2.

Le puits numéro 8 contient un mélange de  
marqueurs de poids moléculaire ADN et les puits 1 à 7  
représentent, dans l'ordre, les produits amplifiés à  
15 partir des ARN totaux de plasmas provenant de 4 témoins  
sains exempts de SEP (puits 1 à 4) et de 3 patients  
atteints de SEP, à différents stades de la maladie  
(puits 5 à 7).

Dans cette série, du matériel nucléaire MSRV-2  
20 est détecté dans le plasma d'un cas de SEP sur les 3  
testés et dans aucun des 4 plasmas témoins. D'autres  
résultats obtenus sur des séries plus étendues confirment  
ces résultats.

La photographie du bas (figure 11) représente le  
25 résultat de l'amplification spécifique par RT-PCR "nichée"  
MSRV-1:

le puits n°1 contient le produit PCR effectué  
avec de l'eau seule, sans addition de transcriptase  
inverse AMV; le puits n°2 contient le produit PCR effectué  
30 avec de l'eau seule, avec addition de transcriptase  
inverse AMV; le puits numéro 3 contient un mélange de  
marqueurs de poids moléculaire ADN; les puits 4 à 13  
contiennent, dans l'ordre, les produits amplifiés à partir  
des ARN totaux extraits de fractions de gradient de  
35 saccharose (collectées du haut vers le bas) sur lequel un  
culot de virion provenant d'un surnageant de culture

infectée par MSRV-1 et MSRV-2 a été centrifugé à l'équilibre selon le protocole décrit par H. Perron (13); dans le puits 14 rien n'a été déposé; dans les puits 15 à 17, on a déposé les produits amplifiés d'ARN extrait de  
5 plasmas provenant de 3 patients différents atteints de SEP, à différents stades de la maladie.

Le génome rétroviral MSRV-1 est bien retrouvé dans la fraction du gradient de saccharose contenant le pic d'activité transcriptase inverse mesurée selon la  
10 technique décrite par H. Perron (3), avec une très forte intensité (fraction 5 du gradient, déposée dans le puits n°8). Une légère amplification a eu lieu dans la première fraction (puits n°4) correspondant vraisemblablement à de l'ARN libéré par des particules  
15 lysées qui flottaient à la surface du gradient; de même, des débris aggrégés ont sédimenté dans la dernière fraction (fond de tube) entraînant quelques copies de génome MSRV-1 qui ont donné lieu à une amplification de faible intensité.

20 Sur les 3 plasmas de SEP testés dans cette série, l'ARN MSRV-1 a été retrouvé dans un cas, produisant une amplification très intense (puits n°17).

Dans cette série, le génome ARN rétroviral MSRV-1 correspondant vraisemblablement à des particules de virus  
25 extracellulaire présentes dans le plasma en nombre extrêmement faible, a été détecté par RT-PCR "nichée" dans un cas de SEP sur les 3 testés. D'autres résultats obtenus sur des séries plus étendues confirment ces résultats.

De plus, la spécificité des séquences amplifiées  
30 par ces techniques PCR peut être vérifiée et évaluée par la technique "ELOS" telle que décrite par F. Mallet (21) et dans le document FR-A-2 663 040.

Pour MSRV-1 les produits de la PCR nichée sus-  
décrite peuvent être testés dans deux systèmes ELOS  
35 permettant de détecter séparément un consensus A et un consensus B+C+D de MSRV-1, correspondant aux sous-familles



décrites dans l'exemple 1 et les figures 1 et 2. En effet, les séquences proches du consensus B+C+D sont retrouvées essentiellement dans les échantillons d'ARN provenant de virions MSRV-1 purifiés à partir de cultures ou amplifiés  
5 dans des liquides biologiques extra-cellulaires de patients SEP, alors que les séquences proches du consensus A sont essentiellement retrouvées dans l'ADN cellulaire humain normal.

Le système ELOSA/MSRV-1 pour la capture et  
10 l'hybridation spécifique des produits PCR de la sous-famille A utilise un oligonucléotide de capture cpV1A avec une liaison amine en 5' et un oligonucléotide de détection biotinylé dpV1A ayant respectivement pour séquence:

- cpV1A identifié par SEQ ID NO 31

15 5' GATCTAGGCCACTTCTCAGGTCCAGS 3', correspondant à l'oligonucléotide de capture ELOSA des produits PCR-nichée MSRV-1 effectuée avec les amorces identifiées par SEQ ID NO 16 et SEQ ID NO 17, suivie éventuellement de l'amplification avec les amorces identifiées par  
20 SEQ ID NO 18 et SEQ ID NO 19 sur prélèvement de patients ;

- dpV1A identifié par SEQ ID NO 32

5' CATCTITTTGGICAGGCAITAGC 3' correspondant à l'oligonucléotide de capture ELOSA de la sous-famille A des produits PCR-"nichée" MSRV-1 effectuée avec les  
25 amorces identifiées par SEQ ID NO 16 et SEQ ID NO 17, suivie éventuellement de l'amplification avec les amorces identifiées par SEQ ID NO 18 et SEQ ID NO 19 sur prélèvement de patients.

Le système ELOSA/MSRV-1 pour la capture et  
30 l'hybridation spécifique des produits PCR de la sous-famille B+C+D utilise le même oligonucléotide de détection biotinylé dpV1A et un oligonucléotide de capture cpV1B avec une liaison amine en 5' ayant pour séquence :

- dpV1B identifié par SEQ ID N°33

35 5' CTTGAGCCAGTTCTCATACCTGGA 3', correspondant à l'oligonucléotide de capture ELOSA de la sous-famille

B + C + D des produits PCR-"nichée" MSRV-1 effectuée avec les amorces identifiées par SEQ ID NO 16 et SEQ ID NO 17, suivie éventuellement de l'amplification avec les amorces identifiées par SEQ ID NO 18 et SEQ ID NO 19 sur  
5 prélèvement de patients.

Ce système de détection ELOSA a permis de vérifier qu'aucun des produits PCR ainsi amplifiés à partir de plasmas de patients SEP traités par la DNase ne contenait de séquence de la sous-famille A et que tous  
10 étaient positifs avec le consensus des sous-familles B, C et D.

Pour MSRV-2, une technique ELOSA similaire a été évaluée sur des isolats provenant de cultures cellulaires infectées en utilisant les amorces d'amplification PCR  
15 suivantes :

- amorce 5', identifiée par SEQ ID NO 34

5' AGTGYTRCCMCARGGCGCTGAA 3', correspondant à une amorce PCR MSRV-2 5', pour PCR sur prélèvement de cultures,

20 - amorce 3', identifiée par SEQ ID NO 35

5' GMGGCCAGCAGSAKGTCATCCA 3', correspondant à une amorce PCR MSRV-2 3', pour PCR sur prélèvement de cultures,

et les oligonucléotides de capture avec une  
25 liaison amine en 5' cpV2, et de détection biotinylé dpV2, ayant pour séquences respectives :

-cpV2 identifiée par SEQ ID NO 36

5 GGATGCCGCCTATAGCCTCTAC 3', correspondant à un oligonucléotide de capture ELOSA des produits de la PCR  
30 MSRV-2 effectuée avec les amorces SEQ ID NO 34 et SEQ ID NO 35, ou éventuellement avec les amorces dégénérées définies par Shih (12),

-dpV2 identifié par SEQ ID NO 37

5' AAGCCTATCGCGTGAGTTGCC 3', correspondant à un  
35 oligonucléotide de détection ELOSA des produits de la PCR MSRV-2 effectuée avec les amorces SEQ ID NO 34 et

SEQ ID NO 35, ou éventuellement avec les amorces dégénérées définies par Shih (12).

5 Ce système d'amplification PCR avec un couple d'amorces différent de ceux qui ont été précédemment décrits pour l'amplification sur les échantillons de patients, a permis de confirmer l'infection par MSRV-2 de cultures in vitro et d'échantillons d'acides nucléiques utilisés pour les travaux de biologie moléculaire.

10 En fin de compte, les premiers résultats de détection PCR du génome d'agents pathogènes et/ou infectants, montrent qu'il est vraisemblable que du "virus" libre peut circuler dans le flux sanguin de patients en phase de poussée aiguë, en dehors du système nerveux. Ceci est compatible avec l'existence  
15 quasi-systématique de "brèches" dans la barrière hémato-encéphalique des patients en phase active de SEP.

**EXEMPLE 7 : OBTENTION DE SEQUENCES DU GENE "env"**  
20 **DU GENOME RETROVIRAL MSRV-1.**

Ainsi que cela a déjà été décrit dans l'exemple 5, une technique PCR dérivée de la technique publiée par Frohman (19) a été utilisée. La technique  
25 dérivée permet, à l'aide d'une amorce spécifique en 3' du génome à amplifier, d'élonguer la séquence vers la région 5' du génome à analyser. Cette variante technique est décrite dans la documentation de la firme "Clontech Laboratories Inc., (Palo-Alto California, USA) fournie  
30 avec son produit "5'-AmplIFINDER™ RACE Kit", qui a été utilisé sur une fraction de virion purifié comme décrit précédemment.

Afin de réaliser une amplification de la région 3' du génome rétroviral MSRV-1 en englobant la  
35 région du gène "env", une étude a été réalisée pour déterminer une séquence consensus dans les régions LTR de

même type que celles du rétrovirus endogène défectif HSERV-9 (18,24), avec lequel le rétrovirus MSRV-1 présente des homologies partielles.

La même amorce 3' spécifique a été utilisée dans le protocole du kit pour la synthèse du ADNc et l'amplification PCR ; sa séquence est la suivante :

GTGCTGATTGGTGTATTTACAATCC (SEQ ID NO 45)

La synthèse de l'ADN complémentaire (ADNc) et l'amplification PCR monodirectionnelle avec l'amorce ci-dessus, ont été réalisées en une étape, selon le procédé décrit dans le brevet EP-A-0 569 272.

Les produits issus de la PCR ont été extraits après purification sur gel d'agarose selon les méthodes conventionnelles (17), puis resuspendus dans 10 ml d'eau distillée. Une des propriétés de la polymérase Taq consistant à ajouter une adénine à l'extrémité 3' de chacun des deux brins d'ADN, l'ADN obtenu a été directement inséré dans un plamide à l'aide du kit TA Cloning<sup>TM</sup> (British Biotechnology). Les 2 µl de solution d'ADN ont été mélangés avec 5 µl d'eau distillée stérile, 1 µl d'un tampon de ligation 10 fois concentré "10X LIGATION BUFFER", 2 µl de "PCR<sup>TM</sup> VECTOR" (25 ng/ml) et 1 µl de "TA DNA LIGASE". Ce mélange a été incubé une nuit à 12°C. Les étapes suivantes ont été réalisées conformément aux instructions du kit TA Cloning<sup>®</sup> (British Biotechnology). A la fin de la procédure, les colonies blanches de bactéries recombinantes (white) ont été repiquées pour être cultivées et permettre l'extraction des plasmides incorporés selon la procédure dite de "miniprep" (17). La préparation de plasmide de chaque colonie recombinante a été coupée par une enzyme de restriction appropriée et analysée sur gel d'agarose. Les plasmides possédant un insert détecté sous lumière UV après marquage du gel au bromure d'éthidium, ont été

sélectionnés pour le séquençage de l'insert, après hybridation avec une amorce complémentaire du promoteur Sp6 présent sur le plasmide de clonage du TA Cloning Kit®. La réaction préalable au séquençage a ensuite été effectuée selon la méthode préconisée pour l'utilisation du kit de séquençage "Prism ready reaction kit dye deoxyterminator cycle sequencing kit" (Applied Biosystems, ref. 401384) et le séquençage automatique a été réalisé avec l'appareil "automatic sequencer, modèle 373 A Applied Biosystems selon les instructions du fabricant.

Cette approche technique a été appliquée à un échantillon de virion concentré comme décrit ci-après, à partir d'un mélange de surnageants de culture produits par des lignées lymphoblastoïdes B telles que décrites dans l'exemple 2, établies à partir des lymphocytes de patients atteints de SEP et présentant une activité transcriptase inverse détectable selon la technique décrite par Perron et coll. (3): les surnageants de cultures sont collectés deux fois par semaine, pré-centrifugés à 10 000 trs/min pendant 30 minutes pour éliminer les débris cellulaires, et ensuite, congelés à -80°C ou utilisés tels quels pour les étapes suivantes. Les surnageants frais ou décongelés sont centrifugés sur un coussin de PBS-glycérol 30% à 100 000g pendant 2 h à 4°C. Après élimination du surnageant, le culot sédimenté constitue l'échantillon de virions concentrés, mais non purifiés. Le culot ainsi obtenu est alors repris dans un petit volume d'un tampon adéquat pour l'extraction d'ARN. La réaction de synthèse de ADNc mentionnée plus haut est réalisée sur cet ARN extrait de virion extracellulaire concentré.

L'amplification RT-PCR selon la technique citée plus-haut a permis d'obtenir le clone FBd3 dont la séquence identifiée par SEQ ID NO 46, est présentée dans la figure 13.

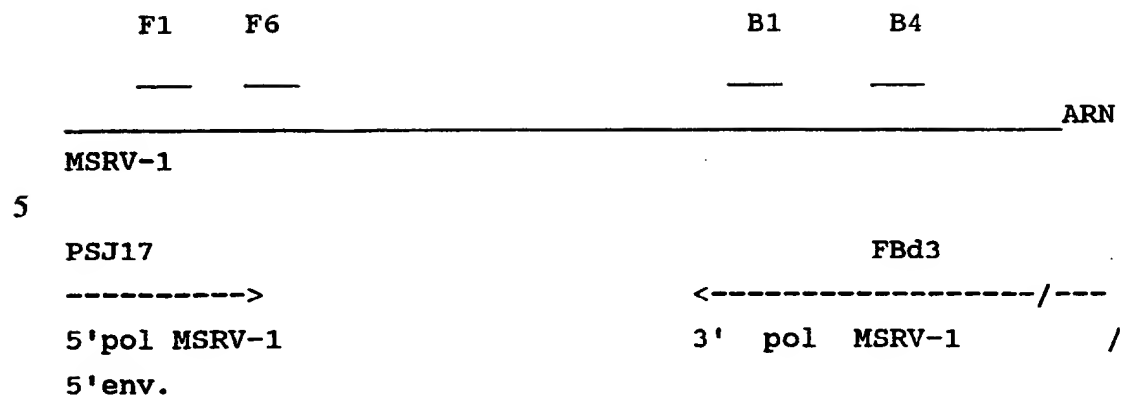
Dans la figure 14, l'homologie de séquence entre le clone FBd3 et le rétrovirus HSERV-9 est représentée sur

le tableau matriciel par un trait plein, pour toute homologie partielle supérieure ou égale à 65 %. On peut constater que des homologies existent dans les régions flanquantes du clone (avec le gène pol en 5' et avec le  
5 gène env puis le LTR en 3'), mais que la région interne est totalement divergente et ne présente aucune homologie, même faible, avec le gène "env" d'HSERV9. De plus, il apparaît que le clone FBd3 contient une région "env" plus longue que celle qui est décrite pour l'endogène défectif  
10 HSERV-9 ; on peut ainsi constater que la région divergente interne constitue un "insert" entre les régions d'homologie partielle avec les gènes défectifs HSERV-9.

15           **EXEMPLE 8 : AMPLIFICATION, CLONAGE ET SEQUENCAGE**  
**DE LA REGION DU GENOME RETROVIRAL MSRV-1 SITUEE ENTRE LES**  
**CLONES PSJ17 ET FBd3.**

Quatre oligonucléotides, F1, B4, F6 et B1, ont  
20 été définis pour amplifier de l'ARN provenant de virions concentrés des souches POL2 et MS7PG. Des réactions de contrôle ont été effectuées de manière à contrôler la présence de contaminants (réaction sur de l'eau). L'amplification consiste en une première étape de RT-PCR  
25 selon le protocole décrit dans la demande de brevet EP-A-0 569 272, suivie d'une deuxième étape de PCR effectuée sur 10 µl de produit de la première étape avec des amorces internes à la première région amplifiée (PCR "nichée"). Dans le premier cycle RT-PCR, les amorces  
30 F1 et B4 sont utilisées. Dans le deuxième cycle PCR, les amorces F6 et l'amorce B1 sont utilisées. Les amorces sont positionnées comme suit :

48



10

Leur composition est :

- amorce F1 : TGATGTGAACGGGCATACTCACTG (SEQ ID NO 47)  
 amorce B4 : CCCAGAGGTTAGGAACTCCCTTTC (SEQ ID NO 48)  
 15 amorce F6 : GCTAAAGGAGACTTGTGGTTGTCAG (SEQ ID NO 49)  
 amorce B1 : CAACATGGGCATTTCGGATTAG (SEQ ID NO 50)

Le produit d'amplification "nichée" obtenu et dénommé "t pol" est présenté dans la figure 15, et correspond à la séquence SEQ ID NO 51.

20

**EXEMPLE 9 : OBTENTION DE NOUVELLES SEQUENCES, EXPRIMEES EN ARN DANS LES CELLULES EN CULTURE PRODUISANT MSRV-1, ET COMPRENANT UNE REGION "env" DU GENOME RETROVIRAL MSRV-1.**

Une banque d'ADNc a été réalisée selon la procédure décrite par le fabricant des kits "cDNA synthesis module, cDNA rapid adaptator ligation module, cDNA rapid cloning module, lambda gt10 in vitro packaging module" (Amersham, ref RPN1256Y/Z, RPN1712, RPN1713, RPN1717, N334Z) à partir de l'ARN messenger extrait de cellules d'une lignée lymphoblastoïde B telle que décrite dans l'exemple 2, établie à partir des lymphocytes d'un patient atteint de SEP et présentant une

activité transcriptase inverse détectable selon la technique décrite par Perron et coll. (3).

Des oligonucléotides ont été définis pour amplifier de l'ADNc cloné dans la banque nucléaire entre la région 3' du clone PSJ17 (pol) et la région 5' (LTR) du clone FBd3. Des réactions de contrôle ont été effectuées de manière à contrôler la présence de contaminants (réaction sur de l'eau). Des réactions PCR effectuées sur les acides nucléiques clonés dans la banque avec différents couples d'amorces ont permis d'amplifier une série clones reliant des séquences pol aux séquences env ou LTR de type MSRV-1.

Deux clones sont représentatifs des séquences obtenues dans la banque de cDNA cellulaire :

- 15 - le clone JLBc1, dont la séquence SEQ ID NO 52 est présentée dans la figure 16 ;
- le clone JLBc2, dont la séquence SEQ ID NO 53 est présentée dans la figure 17.

Les séquences des clones JLBc1 et JLBc2 sont homologues à celle du clone FBd3, ainsi que cela apparaît dans les figures 18 et 19. L'homologie entre le clone JLBc1 et le clone JLBc2 est représentée dans la figure 20.

Les homologies entre les clones JLBc1 et JLBc2 d'une part, et la séquence HSERV9 d'autre part, sont présentées respectivement dans les figures 21 et 22.

On remarque que la région d'homologie entre JLB1, JLB2 et FBd3 comprend, avec quelques variations de séquence et de taille de "l'insert", la séquence supplémentaire absente ("insérée") dans la séquence env HSERV-9, telle que décrite dans l'exemple 8.

On remarque aussi que la région "pol" clonée est très homologue à HSERV-9, ne possède pas de cadre de lecture (en tenant compte des erreurs de séquences induites par les techniques utilisées, jusqu'au séquenceur automatique inclus) et diverge des séquences MSRV-1 obtenues à partir de virions. Etant donné que ces



séquences ont été clonées à partir de l'ARN de cellules exprimant des particules MSRV-1, il est probable qu'elles proviennent d'éléments rétroviraux endogènes apparentés à la famille ERV9 ; ce, d'autant plus que les gènes pol et env sont présent sur un même ARN qui n'est à l'évidence pas l'ARN génomique MSRV-1. Certains de ces éléments ERV9 possèdent des LTR fonctionnels activables par des virus répliatifs codant pour des transactivateurs homologues ou hétérologues. Dans ces conditions, la parenté entre MSRV-1 et HSERV-9 rend probable la transactivation des éléments ERV9 endogènes défectifs (ou non) par des protéines transactivatrices MSRV-1 homologues, voire identiques.

Un tel phénomène peut induire une interférence virale entre l'expression de MSRV-1 et les éléments endogènes apparentés. Une telle interférence conduit généralement à une expression dite "défective-interférente", dont certaines caractéristiques ont été retrouvées dans les cultures étudiées infectées par MSRV-1. De plus, un tel phénomène n'est pas sans générer l'expression de polypeptides, voire de protéines rétrovirales endogènes qui ne sont pas nécessairement tolérées par le système immunitaire. Un tel schéma d'expression aberrante d'éléments endogènes apparentés à MSRV-1 et induite par ce dernier est susceptible de multiplier les antigènes aberrants et, donc, de contribuer à l'induction de processus autoimmuns tels qu'on les observe dans la SEP.

Il est cependant essentiel de noter que les clones JLBc1 et JLBc2 diffèrent de la séquence ERV9 ou HSERV9 déjà décrite, en ce qu'ils possèdent une région env plus longue comprenant une région supplémentaire totalement divergente d'ERV9. On peut donc définir leur parenté avec la famille endogène ERV9, mais ils constituent à l'évidence des éléments originaux, jamais décrits à ce jour. En effet, l'interrogation des banques de données de séquences nucléiques disponibles dans la

version n° 15 (1995) du logiciel "Entrez" (NCBI, NIH, Bethesda, USA) n'a pas permis d'identifier de séquence homologue connue dans la région env de ces clones.

5

**EXEMPLE 10 : OBTENTION DES SEQUENCES SITUEES DANS LA REGION 5' pol et 3' gag DU GENOME RETROVIRAL MSRV-1.**

Ainsi que cela a déjà été décrit dans l'exemple 5, une technique PCR dérivée de la technique publiée par Frohman (19) a été utilisée. La technique dérivée permet, à l'aide d'une amorce spécifique en 3' du génome à amplifier, d'élonguer la séquence vers la région 5' du génome à analyser. Cette variante technique est décrite dans la documentation de la firme "Clontech Laboratories Inc., (Palo-Alto California, USA) fournie avec son produit "5'-AmpliFINDER™ RACE Kit", qui a été utilisé sur une fraction de virion purifié comme décrit précédemment.

Afin de réaliser une amplification de la région 5' du génome rétroviral MSRV-1 partant de la séquence pol déjà séquencée (clone F11-1) et s'étendant vers le gène gag, des amorces spécifiques MSRV-1 ont été définies.

Les amorces 3' spécifiques utilisées dans le protocole du kit pour la synthèse du ADNc et l'amplification PCR sont respectivement complémentaires aux séquences MSRV-1 suivantes :

ADNc : (SEQ ID NO 54)  
CCTGAGTTCTTGCACTAACCC  
amplification : (SEQ ID NO 55)  
GTCCGTTGGGTTTCCTTACTCCT

Les produits issus de la PCR ont été extraits après purification sur gel d'agarose selon les méthodes

conventionnelles (17), puis resuspendus dans 10 ml d'eau distillée. Une des propriétés de la polymérase Taq consistant à ajouter une adénine à l'extrémité 3' de chacun des deux brins d'ADN, l'ADN obtenu a été directement  
5 inséré dans un plamide à l'aide du kit TA Cloning™ (British Biotechnology). Les 2 µl de solution d'ADN ont été mélangés avec 5 µl d'eau distillée stérile, 1 µl d'un tampon de ligation 10 fois concentré "10X LIGATION BUFFER", 2 µl de "pCR™ VECTOR" (25 ng/ml)  
10 et 1 µl de "TA DNA LIGASE". Ce mélange a été incubé une nuit à 12°C. Les étapes suivantes ont été réalisées conformément aux instructions du kit TA Cloning® (British Biotechnology). A la fin de la procédure, les colonies blanches de bactéries recombinantes (white) ont  
15 été repiquées pour être cultivées et permettre l'extraction des plasmides incorporés selon la procédure dite de "miniprep" (17). La préparation de plasmide de chaque colonie recombinante a été coupée par une enzyme de restriction appropriée et analysée sur gel d'agarose. Les  
20 plasmides possédant un insert détecté sous lumière UV après marquage du gel au bromure d'éthidium, ont été sélectionnés pour le séquençage de l'insert, après hybridation avec une amorce complémentaire du promoteur Sp6 présent sur le plasmide de clonage du TA Cloning Kit®.  
25 La réaction préalable au séquençage a ensuite été effectuée selon la méthode préconisée pour l'utilisation du kit de séquençage "Prism ready reaction kit dye deoxyterminator cycle sequencing kit" (Applied Biosystems, ref. 401384) et le séquençage automatique a été réalisé  
30 avec l'appareil "automatic sequencer, modèle 373 A Applied Biosystems selon les instructions du fabricant.

Cette approche technique a été appliquée à un échantillon de virion concentré comme décrit ci-après, à partir d'un mélange de surnageants de culture produits par  
35 des lignées lymphoblastoïdes B telles que décrites dans l'exemple 2, établies à partir des lymphocytes de patients

atteints de SEP et présentant une activité transcriptase inverse détectable selon la technique décrite par Perron et coll. (3) : les surnageants de cultures sont collectés deux fois par semaine, pré-centrifugés à 10 000 trs/min pendant 30 minutes pour éliminer les débris cellulaires, et ensuite, congelés à -80°C ou utilisés tels quels pour les étapes suivantes. Les surnageants frais ou décongelés sont centrifugés sur un coussin de PBS-glycérol 30 % à 100 000g pendant 2 h à 4°C. Après élimination du surnageant, le culot sédimenté constitue l'échantillon de virions concentrés, mais non purifiés. Le culot ainsi obtenu est alors repris dans un petit volume d'un tampon adéquat pour l'extraction d'ARN. La réaction de synthèse de ADNC mentionnée plus haut est réalisée sur cet ARN extrait de virion extracellulaire concentré.

L'amplification RT-PCR selon la technique citée plus-haut a permis d'obtenir le clone GM3 dont la séquence identifiée par SEQ ID NO 56, est présentée dans la figure 23.

Dans la figure 24, l'homologie de séquence entre le clone GMP3 et le rétrovirus HSERV-9 est représentée sur le tableau matriciel par un trait plein, pour toute homologie partielle supérieure ou égale à 65 %.

En synthèse, sur la figure 25 est représentée la localisation des différents clones précédemment étudiés, par rapport au génome ERV9 connu. Sur la figure 25, la région env MSRV-1 étant plus longue que le gène env de ERV9 de référence, la région supplémentaire est représentée au-dessus du point d'insertion selon un "V", étant entendu que le matériel inséré présente une variabilité de séquences et de taille entre les clones représentés (JLBc1, JLBc2, FBd3). Et sur la figure 26, est représentée la position de différents clones étudiés dans la région MSRV-1 pol \*.

Grâce au clone GM3 précédemment décrit, on a pu définir un cadre de lecture possible, couvrant l'ensemble

du gène pol, référencé selon SEQ ID NO 57, représenté aux figures successives 27a à 27c.

5           **EXEMPLE 11 : DETECTION D'ANTICORPS SPECIFIQUES ANTI-MSRV-1 DANS LE SERUM HUMAIN.**

L'identification de la séquence du gène pol du rétrovirus MSRV-1 et d'un cadre de lecture ouverte de ce  
10 gène a permis de déterminer la séquence SEQ ID NO 39 en acides aminés d'une région dudit gène, référencée SEQ ID NO 40 (cf figure 28).

Différents peptides synthétiques correspondant à des fragments de la séquence protéique de la transcriptase  
15 inverse MSRV-1 codée par le gène pol, ont été testés pour leur spécificité antigénique vis à vis de sera de patients atteints de SEP et de témoins sains.

Les peptides ont été synthétisés chimiquement par synthèse en phase solide, selon la technique de Merrifield  
20 (Barany G, and Merrifield R.B, 1980, In the Peptides, 2, 1-284, Gross E and Meienhofer J, Eds Academic Press, New York). Les modalités pratiques sont celles décrites ci-après.

a) Synthèse des peptides:

25 Les peptides ont été synthétisés sur une résine phénylacétamidométhyle (PAM) /polystyrène/divinylbenzène (Applied Biosystems, Inc. Foster City, CA), en utilisant un synthétiseur automatique "Applied Biosystems 430A". Les acides aminés sont couplés sous forme d'esters  
30 d'hydroxybenzotriazole (HOBt). Les acides aminés utilisés proviennent de Novabiochem (Laüflerlfingen, Suisse) ou de Bachem (Bubendorf, Suisse).

La synthèse chimique a été effectuée en utilisant un protocole de double couplage avec la N-méthyl-pyrrolidone (NMP) comme solvant. Les peptides ont été  
35 coupés de la résine ainsi que les protections latérales,

de manière simultanée, à l'aide d'acide fluorhydrique (HF) dans un appareil approprié (appareil de coupure de type I, Peptide Institute, Osaka, Japon).

Pour 1g de peptidylrésine, 10ml de HF, 1ml  
5 d'anisole et 1ml de diméthylsulfure 5DMS sont utilisés. Le mélange est agité durant 45 minutes à -2°C. Le HF est alors évaporé sous vide. Après lavages intensifs à l'éther, le peptide est élué de la résine par de l'acide acétique 10%, puis lyophilisé.

10 Les peptides sont purifiés par chromatographie liquide haute performance préparative sur une colonne VYDAC de type C18 (250 x 21 mm) (The Separation Group, Hesperia, CA, USA). L'élution est réalisée par un gradient  
15 d'acétonitrile à un débit de 22 ml/min. Les fractions collectées sont contrôlées par une élution en condition isocratique sur une colonne VYDAC®. C18 analytique (250 x 4,6 mm), à un débit de 1ml/min. Les fractions qui  
20 présentent le même temps de rétention sont réunies et lyophilisées. La fraction majoritaire est ensuite analysée par chromatographie liquide haute performance analytique, avec le système décrit précédemment. Le peptide qui est considéré comme de pureté acceptable se traduit par un pic unique représentant 95 % minimum du chromatogramme.

Les peptides purifiés sont ensuite analysés dans  
25 le but de contrôler leur composition en acides aminés, à l'aide d'un analyseur d'acides aminés automatique Applied Biosystems 420H. La mesure de la masse moléculaire chimique (moyenne) des peptides est obtenue en utilisant la spectrométrie de masse L.S.I.M.S. en mode d'ion positif  
30 sur un instrument à double focalisation VG. ZAB.ZSEQ relié à un système d'acquisition DEC-VAX 2000 (VG analytical Ltd, Manchester, Angleterre).

La réactivité des différents peptides a été testée contre des sera de patients atteints de SEP et  
35 contre des sera de témoins sains. Ceci a permis de sélectionner un peptide dénommé POL2B dont la séquence est

représentée à la figure 28 dans l'identificateur SEQ ID NO 39 ci-dessous, codé par le gène *pol* de MSRV-1 (nucléotides 181 à 330).

b) Propriétés antigéniques :

- 5 Les propriétés antigéniques du peptide POL2B ont été mises en évidence selon le protocole ELISA décrit ci-dessous.

- Le peptide POL2B lyophilisé a été dissous dans de l'eau distillée stérile à une concentration de 1mg/ml.
- 10 Cette solution-mère a été aliquotée et gardée à +4°C pour usage sous quinzaine, ou congelée à -20°C, pour un usage dans les 2 mois. Un aliquot est dilué dans une solution de PBS (Phosphate Buffer Saline) afin d'obtenir une concentration finale de peptide de 1 microgramme/ml.
- 15 100 microlitres de cette dilution sont déposés dans chaque puits de plaques de microtitration (plastique "high-binding", COSTAR ref: 3590). Les plaques sont recouvertes d'un adhésif de type "plate-sealer" et maintenues une nuit à +4°C, pour la phase d'adsorption du
- 20 peptide sur le plastique. L'adhésif est enlevé et les plaques sont lavées trois fois avec un volume de 300 microlitres d'une solution A (PBS 1x, 0,05% Tween 20®), puis retournées sur un tissu absorbant. Les plaques ainsi égouttées sont remplies avec 200 microlitres par puits
- 25 d'une solution B (solution A + 10 % de sérum de chèvre), puis recouvertes d'un adhésif et incubées de 45 minutes à 1 heure, à 37°C. Les plaques sont ensuite lavées trois fois avec la solution A, comme décrit précédemment.

- Les échantillons de sérum à tester sont
- 30 préalablement dilués au 1/50ème dans la solution B et 100 microlitres de chaque sérum dilué à tester sont déposés dans les puits de chaque plaque de microtitration. Un contrôle négatif est déposé dans un puits de chaque plaque, sous la forme de 100 microlitres de tampon B. Les
- 35 plaques recouvertes d'un adhésif sont alors incubées 1 à 3 heures, à 37°C. Les plaques sont ensuite lavées trois

fois avec la solution A, comme décrit précédemment. Parallèlement, un anticorps de chèvre marqué à la peroxydase et dirigé contre les IgG (Sigma Immunochemicals ref. A6029) ou les IgM (Cappel ref.: 55228) humaines est  
5 dilué dans la solution B (dilution 1/5000 pour l'anti-IgG et, 1/1000 pour l'anti-IgM). 100 microlitres de la dilution adéquate de l'anticorps marqué sont alors déposés dans chaque puits des plaques de microtitration et les plaques recouvertes d'un adhésif sont incubées 1 à  
10 2 heures, à 37°C. Un nouveau lavage des plaques est ensuite effectué comme décrit précédemment. Parallèlement, le substrat de la peroxydase est préparé selon les indications du kit "Sigma fast OPD kit" (Sigma Immunochemichals, ref. P9187). 100 microlitres de  
15 solution-substrat sont déposés dans chaque puits et les plaques sont placées à l'abri de la lumière pendant 20 à 30 minutes, à température ambiante.

Une fois la réaction colorée stabilisée, les plaques sont immédiatement placées dans un lecteur  
20 spectrophotométrique de plaques ELISA, et la densité optique (D.O.) de chaque puits est lue à une longueur d'onde de 492 nm. Alternativement, 30 microlitres d'HCL 1N sont déposés dans chaque puits pour stopper la réaction et les plaques sont lues au spectrophotomètre dans les  
25 24 heures.

Les échantillons sérologiques sont déposés en double ou en triple et la densité optique (D.O.) correspondant au sérum testé est calculée en faisant la moyenne des D.O. obtenues pour un même échantillon, à la  
30 même dilution.

La D.O. nette de chaque sérum correspond à la D.O. moyenne du sérum de laquelle est soustraite la D.O. moyenne du témoin négatif (solution B : PBS, Tween 20® 0,05%, 10% sérum de chèvre).

35 c) Détection d'anticorps IgG anti-MSRV-1 par ELISA :



La technique décrite ci-dessus a été utilisée avec le peptide POLB2 pour rechercher la présence d'anticorps IgG spécifiques anti-MSRV-1 dans le sérum de 29 patients, pour lesquels un diagnostic certain ou probable de SEP a été posé selon les critères de Poser (23), et de 32 témoins sains (donneurs de sang).

Dans la figure 29, les résultats de chaque sérum testé avec un anticorps anti-IgG, sont représentés. Chaque barre verticale représente la densité optique (D.O. à 492 nm) nette d'un sérum testé. L'axe des ordonnées indique la D.O. nette au sommet des barres verticales. Les 29 premières barres verticales situées à gauche de la ligne pointillée verticale, représentent les sera de 29 cas de SEP testés et les 32 barres verticales situées à droite de la ligne pointillée verticale, représentent les sera de 32 témoins sains (donneurs de sang).

La moyenne des D.O. nettes des SEP testées est de : 0,62. Le graphique permet de visualiser 5 témoins dont la D.O. nette émerge au dessus des valeurs groupées de la population témoin. Ces valeurs peuvent représenter la présence d'IgG spécifiques chez des patients séropositifs non-malades. Deux méthodes ont donc été évaluées pour déterminer le seuil statistique de positivité du test.

La moyenne des D.O. nettes des témoins, y compris les témoins avec des D.O. nettes élevées, est de 0,36. Sans les 5 témoins dont les D.O. nettes sont supérieures ou égales à 0,5, la moyenne des témoins "négatifs" est de 0,33. L'écart type des témoins négatifs est de 0,10. Un seuil de positivité théorique peut être calculé selon la formule :

valeur seuil (moyenne des D.O. nettes des témoins séronégatifs) + (2 ou 3 x écart-type des D.O. nettes des témoins séronégatifs).

Dans le premier cas, on considère qu'il existe des séropositifs non-malades et la valeur seuil est égale à :  $0,33 + (2 \times 0,10) = 0,53$ . Les résultats négatifs représentent un "bruit de fond" non spécifique de la présence d'anticorps spécifiquement dirigés contre un épitope du peptide.

Dans le deuxième cas, si l'ensemble des témoins constitués de donneurs de sang en bonne santé apparente est pris comme base de référence, sans exclure les sérums a priori séropositifs, l'écart-type des "témoins non-SEP" est de 0,116. La valeur seuil devient alors :  $0,36 + (2 \times 0,116) = 0,59$ .

Selon cette analyse, le test est spécifique de la SEP. A cet égard, on constate que le test est spécifique de la SEP, puisque comme montré dans le tableau 1, aucun témoin n'a une D.O. nette au-dessus de ce seuil. En fait ce résultat reflète le fait que les titres d'anticorps chez les patients atteints de SEP sont en majorité plus élevés que chez des témoins sains ayant été en contact avec MSRV-1.

TABLEAU N°1

	SEP	TEMOINS
	0,681	0,3515
	1,0425	0,56
	0,5675	0,3565
	0,63	0,449
	0,588	0,2825
	0,645	0,55
	0,6635	0,52
	0,576	0,2535
	0,7765	0,55
	0,5745	0,51
	0,513	0,426
	0,4325	0,451
	0,7255	0,227
	0,859	0,3905
	0,6435	0,265
	0,5795	0,4295
	0,8655	0,291
	0,671	0,347
	0,596	0,4495
	0,662	0,3725
	0,602	0,181
	0,525	0,2725
	0,53	0,426
	0,565	0,1915
	0,517	0,222
	0,607	0,395
	0,3705	0,34
	0,397	0,307
	0,4395	0,219
		0,491
		0,2265
		0,2605
MOYENNE	0,62	0,33
ECART TYPE	0,14	0,10
VALEUR SEUIL		0,53

En fonction du premier mode de calcul et comme représenté à la figure 29 et dans le tableau correspondant 1, 26 des 29 sérums SEP donnent un résultat positif (D.O. nette supérieure ou égale à 0,50) indiquant la présence d'IgG spécifiquement dirigées contre le peptide POL2B, donc contre une partie de l'enzyme transcriptase inverse du rétrovirus MSRV-1 codée par son gène pol et, par conséquent, contre le rétrovirus MSRV-1. Ainsi, environ 90 % des patients SEP testés ont réagi contre un épitope porté par le peptide POL2B et présentent des IgG circulantes dirigées contre ce dernier.

Cinq donneurs de sang, en apparente bonne santé, sur 32 présentent un résultat positif. Ainsi, il apparaît qu'environ 15 % de la population non-malade peut avoir été en contact d'un épitope porté par le peptide POL2B dans des conditions ayant conduit à une immunisation active qui se traduit par la persistance d'IgG sériques spécifiques. Ces conditions sont compatibles avec une immunisation contre la transcriptase inverse du rétrovirus MSRV-1, lors d'une infection par le (et/ou réactivation du) rétrovirus MSRV-1. L'absence de pathologie neurologique apparente évoquant la SEP chez ces témoins séropositifs peut indiquer qu'ils sont porteurs sains, ont éliminé un virus infectieux après s'être immunisés ou qu'ils constituent une population à risque de porteurs chroniques. En effet, les données épidémiologiques montrant qu'un agent pathogène présent dans l'environnement des régions à haute prévalence de SEP, peut être la cause de cette maladie, impliquent qu'une fraction de la population exempte de SEP a nécessairement été en contact avec un tel agent pathogène. On a montré que le rétrovirus MSRV-1 constitue tout ou partie de cet "agent pathogène" à l'origine de la SEP et il est donc normal que des témoins pris dans une population saine présentent des anticorps de type IgG contre des composants du rétrovirus MSRV-1. Ainsi, la différence de séroprévalence entre les SEP et la

population témoin est extrêmement significative : test "chi-2",  $p < 0,001$ . Ces résultats sont donc en faveur d'un rôle étiopathogénique de MSRV-1 dans la SEP.

d) Détection d'anticorps IgM anti-MSRV-1 par

5 ELISA :

La technique ELISA avec le peptide POL2B a été utilisée pour rechercher la présence d'anticorps spécifiques IgM anti-MSRV-1 dans le sérum de 36 patients, pour lesquels un diagnostic certain ou probable de SEP a  
10 été posé selon les critères de Poser (23), et de 42 témoins sains (donneurs de sang).

Dans la figure 30, les résultats de chaque sérum testé avec un anticorps anti-IgM, sont représentés. Chaque barre verticale représente la densité optique (D.O à  
15 492 nm) nette d'un sérum testé. L'axe des ordonnées indique la D.O. nette au sommet des barres verticales. Les 36 premières barres verticales situées à gauche de la ligne verticale coupant l'axe des abscisses représentent les sérums de 36 cas de SEP testés et les barres  
20 verticales situées à droite de la ligne pointillée verticale, représentent les sérums de 42 témoins sains (donneurs de sang). La ligne horizontale tracée au milieu du graphique représente un seuil théorique délimitant les résultats positifs (dont le sommet de la barre est situé  
25 au dessus) et les résultats négatifs (dont le sommet de la barre est situé au dessous).

La moyenne des D.O. nettes des SEP testées est de : 0,19.

La moyenne des D.O. nettes des témoins est  
30 de : 0,09.

L'écart type des témoins négatifs est de : 0,05.

Etant donné la faible différence entre la moyenne et l'écart-type des témoins, le seuil de positivité théorique peut être calculé selon la formule :

- 5 valeur seuil = (moyenne des D.O. nettes des témoins séronégatifs) + (3 x écart-type des D.O. nettes des témoins séronégatifs).

La valeur seuil est donc égale à  $0,09 + (3 \times 0,05) = 0,26$  ; soit, en pratique, 0,25.

- 10 Les résultats négatifs représentent un "bruit de fond" non spécifique de la présence d'anticorps spécifiquement dirigés contre un épitope du peptide.

- Selon cette analyse et comme montré à la figure 30 et dans le tableau 2 correspondant, le test IgM  
15 est spécifique de la SEP, puisqu'aucun témoin n'a une D.O. nette au-dessus du seuil. 7 des 36 sérums SEP produisent un résultat IgM positif ; or l'étude des données cliniques révèle que ces sérums positifs ont été prélevés lors d'une première poussée de SEP ou d'une poussée aiguë chez des  
20 malades non-traités. Il est connu que les IgM dirigées contre des agents pathogènes sont produites lors des primo-infections ou lors de réactivations suivant une phase de latence dudit agent pathogène.

- La différence de séroprévalence entre les SEP et  
25 la population témoin est extrêmement significative : test "chi-2",  $p < 0,001$ .

Ces résultats sont en faveur d'un rôle étiopathogénique de MSRV-1 dans la SEP.

- La détection des anticorps IgM et IgG contre le  
30 peptide POL2B permet d'évaluer l'évolution d'une infection par MSRV-1 et/ou de la réactivation virale de MSRV-1.

TABLEAU N°2

	SEP	TEMOINS
	0,064	0,243
	0,087	0,11
	0,044	0,098
	0,115	0,028
	0,089	0,094
	0,025	0,038
	0,097	0,176
	0,108	0,146
	0,018	0,049
	0,234	0,161
	0,274	0,113
	0,225	0,079
	0,314	0,093
	0,522	0,127
	0,306	0,02
	0,143	0,052
	0,375	0,062
	0,142	0,074
	0,157	0,043
	0,168	0,046
	1,051	0,041
	0,104	0,13
	0,187	0,153
	0,044	0,107
	0,053	0,178
	0,153	0,114
	0,07	0,078
	0,033	0,118
	0,104	0,177
	0,187	0,026
	0,044	0,024
	0,053	0,046
	0,153	0,116
	0,07	0,04
	0,033	0,028
	0,973	0,073
		0,008
		0,074
		0,141
		0,219
		0,047
		0,017
MOYENNE	0,19	0,09
ECART TYPE	0,23	0,05
VALEUR SEUIL		0,26

e) Recherche d'épitopes immunodominants dans le peptide POL2B :

Afin de réduire le bruit de fond non-spécifique et d'optimiser la détection des réponses des anticorps anti-MSRV-1, la synthèse d'octapeptides, successivement décalés d'un acide aminé, couvrant l'ensemble de la séquence déterminée par POL2B, a été réalisée selon le protocole décrit ci-dessous.

La synthèse chimique d'octapeptides chevauchants couvrant la séquence en acides aminés 61-110 représentée dans l'identificateur SEQ ID NO 39 a été réalisée sur membrane de cellulose activée selon la technique de BERG et al. (1989. J. Ann. Chem. Soc., 111, 8024-8026) commercialisée par Cambridge Research Biochemicals sous la dénomination commerciale Spotscreen. Cette technique permet la synthèse simultanée d'un grand nombre de peptides et leur analyse.

La synthèse est réalisée avec des acides aminés estérifiés dont le groupement  $\alpha$ -aminé est protégé par un groupement FMOC (Nova Biochem) et les groupements latéraux par des groupements protecteurs tels que trityl, t-butyl ester ou t-butyl éther. Les acides aminés estérifiés sont solubilisés dans du N-méthyl pyrrolidone (NMP) à la concentration de 300 nM, et 0,9  $\mu$ l sont déposés au niveau de taches de dépôt de bleu de bromophénol. Après incubation de 15 minutes, un nouveau dépôt d'acides aminés est réalisé suivi d'une autre incubation de 15 minutes. Si le couplage entre deux acides aminés s'est effectué correctement on observe une modification de coloration (passage du bleu au vert-jaune). Après trois lavages dans du DMF, une étape d'acétylation est effectuée par de l'anhydride acétique. Ensuite, les groupements aminés terminaux des peptides en cours de synthèse sont déprotégés par de la pipéridine 20 % dans le DMF. Les taches de dépôt sont recolorées par une solution de bleu de bromophénol à 1 % dans le DMF, lavées trois fois au



méthanol et séchées. L'ensemble de ces opérations constitue un cycle d'addition d'un acide aminé et ce cycle est répété jusqu'à l'achèvement de la synthèse. Lorsque tous les acides aminés ont été ajoutés, le groupement  
5 NH<sub>2</sub>-terminal du dernier acide aminé est déprotégé par de la pipéridine à 20 % dans le DMF et acétylé par de l'anhydride acétique. Les groupements protecteurs de la chaîne latérale sont enlevés par un mélange de dichlorométhane/acide trifluoroacétique/triisobutylsilane  
10 (5ml/5ml/250µl). L'immunoréactivité des peptides est ensuite testée par ELISA.

Après synthèse en double des différents octapeptides sur deux membranes différentes, ces dernières sont rincées avec du méthanol et lavées dans du TBS  
15 (Tris 0,1M pH 7,2), puis incubées une nuit à température ambiante dans un tampon de saturation. Après plusieurs lavages dans du TBS-T (Tris 0,1M pH 7,2 - 0,05% Tween 20), une membrane est incubée avec une dilution au 1/50 d'un sérum de référence provenant d'un patient atteint de SEP  
20 et l'autre membrane avec une dilution au 1/50 d'un pool de sérums de témoins sains. Les membranes sont incubées 4 heures à température ambiante. Après lavages avec du TBS-T, un conjugué anti-immunoglobulines humaines marqué à la  $\beta$ -galactosidase (commercialisé par Cambridge Research  
25 Biomedicals) est ajouté à une dilution au 1/200 et l'ensemble est incubé deux heures à température ambiante. Après lavages des membranes par du TBS-T 0,05% et du PBS, l'immunoréactivité au niveau des différentes taches est révélée par addition de 5-bromo-4-chloro-3-indoyl- $\beta$ -D-  
30 galactopyranoside dans du potassium. L'intensité de coloration des taches est estimée qualitativement avec une valeur relative de 0 à 5 comme représenté aux figures 31 à 33 annexées.

On peut ainsi déterminer deux régions  
35 immunodominantes à chaque extrémité du peptide POL2B correspondant respectivement aux séquences en acides

aminés 65-75 (SEQ ID NO 41) et 92-109 (SEQ ID NO 42), selon figure 34, et comprises respectivement entre les octapeptides Phe-Cys-Ile-Pro-Val-Arg-Pro-Asp (FCIPVRPD) et Arg-Pro-Asp-Ser-Gln-Phe-Leu-Phe (RPDSQFLF), et, Thr-Val-  
5 Leu-Pro-Gln-Gly-Phe-Arg (TVLPQGFR) et Leu-Phe-Gly-Gln-Ala-Leu-Ala-Gln (LFGQALAQ) et une région moins réactive, mais apparemment plus spécifique, puisqu'elle ne produit aucun bruit de fond avec le sérum témoin, représentée par les octapeptides Leu-Phe-Ala-Phe-Glu-Asp-Pro-Leu (LFAFEDPL)  
10 (SEQ ID NO 43) et Phe-Ala-Phe-Glu-Asp-Pro-Leu-Asn (FAFEDPLN) (SEQ ID NO 44).

Ces régions permettent de définir de nouveaux peptides plus spécifiques et plus immunoréactifs selon les techniques habituelles.

15 Il est ainsi possible, grâce aux découvertes effectuées et aux méthodes mises au point par les inventeurs, de réaliser un diagnostic de l'infection et/ou de la réactivation MSRV-1, et d'évaluer une thérapie dans la SEP sur la base de son efficacité à "négativer" la  
20 détection de ces agents dans les fluides biologiques des patients. De plus, la détection précoce chez des personnes ne présentant pas encore de signes neurologiques de SEP, pourrait permettre d'instaurer un traitement d'autant plus efficace sur l'évolution clinique ultérieure qu'il  
25 précèderait le stade lésionnel qui correspond à l'apparition des troubles neurologiques. Or, à ce jour, un diagnostic de SEP ne peut être établi avant l'installation d'une symptomatologie neurologique lésionnelle et, donc, aucun traitement n'est instauré avant l'émergence d'une  
30 clinique évocatrice de lésions du système nerveux central déjà conséquentes. Le diagnostic d'une infection et/ou réactivation de MSRV-1 et/ou MSRV-2 chez l'homme est donc déterminant et la présente invention en fournit les moyens.

35 Il est ainsi possible, outre de réaliser un diagnostic de l'infection et/ou de la réactivation MSRV-1,

d'évaluer une thérapie dans la SEP sur la base de son efficacité à "négativer" la détection de ces agents dans les fluides biologiques des patients.

5                    **EXEMPLE 12 : OBTENTION D'UN CLONE LB19 CONTENANT  
UNE PARTIE DU GENE GAG DU RETROVIRUS MSRV-1**

Une technique PCR dérivée de la technique publiée par Gonzalez-Quintial R et al. (19) et PLAZA et al (25) a  
10 été utilisée. A partir des ARNs totaux extrait d'une fraction de virion purifié comme décrit précédemment, la cDNA a été synthétisé à l'aide d'une amorce spécifique (SEQ ID N°64) en 3' du génome à amplifier, en utilisant la EXPAND<sup>TM</sup> REVERSE TRANSCRIPTASE (BOEHRINGER MANNHEIM).

15

cDNA :

AAGGGGCATG GACGAGGTGG TGGCTTATT (SEQ ID NO°65) (anti-sens)

Après purification, une queue poly G a été ajoutée à  
20 l'extrémité 5' du cDNA à l'aide du kit "Terminal transferases kit" commercialisé par la société Boehringer Mannheim, selon le protocole du fabriquant.

Une PCR d'ancrage a été réalisée à l'aide des amorces  
25 5' et 3'suivantes:

AGATCTGCAG AATTCGATAT CACCCCCCCC CCCCCC (SEQ ID N°91) (sens), et

AAATGTCTGC GGCACCAATC TCCATGTT (SEQ ID NO°64) (anti-sens)

Ensuite, une PCR d'ancrage semi-nichée a été réalisée avec les amorces 5' et 3'suivantes:

30 AGATCTGCAG AATTCGATAT CA (SEQ ID N°92) (sens), et

AAATGTCTGC GGCACCAATC TCCATGTT (SEQ ID NO°64) (anti-sens)

Les produits issus de la PCR ont été purifiés  
35 après purification sur gel d'agarose selon les méthodes conventionnelles (17), puis resuspendus dans

10 microlitres d'eau distillée. Une des propriétés de la polymérase Taq consistant à ajouter une adénine à l'extrémité 3' de chacun des deux brins d'ADN, l'ADN obtenu a été directement inséré dans un plamide à l'aide du kit TA Cloning<sup>TM</sup> (British Biotechnology). Les 2  $\mu$ l de solution d'ADN ont été mélangés avec 5  $\mu$ l d'eau distillée stérile, 1  $\mu$ l d'un tampon de ligation 10 fois concentré "10X LIGATION BUFFER", 2  $\mu$ l de "pCR<sup>TM</sup> VECTOR" (25 ng/ml) et 1  $\mu$ l de "T4 DNA LIGASE". Ce mélange a été incubé une nuit à 12°C. Les étapes suivantes ont été réalisées conformément aux instructions du kit TA Cloning® (British Biotechnology). A la fin de la procédure, les colonies blanches de bactéries recombinantes (white) ont été repiquées pour être cultivées et permettre l'extraction des plasmides incorporés selon la procédure dite de "miniprep" (17). La préparation de plasmide de chaque colonie recombinante a été coupée par une enzyme de restriction appropriée et analysée sur gel d'agarose. Les plasmides possédant un insert détecté sous lumière UV après marquage du gel au bromure d'éthidium, ont été sélectionnés pour le séquençage de l'insert, après hybridation avec une amorce complémentaire du promoteur Sp6 présent sur le plasmide de clonage du TA Cloning Kit®. La réaction préalable au séquençage a ensuite été effectuée selon la méthode préconisée pour l'utilisation du kit de séquençage "Prism ready reaction kit dye deoxyterminator cycle sequencing kit" (Applied Biosystems, ref. 401384) et le séquençage automatique a été réalisé avec l'appareil "Automatic Sequencer, modèle 373 A" Applied Biosystems selon les instructions du fabricant.

L'amplification PCR selon la technique citée plus-haut a été utilisée sur un ADNc synthétisé à partir des acides nucléiques de fractions de particules infectantes purifiées sur gradient de saccharose, selon la technique décrite par H. Perron (13), à partir de surnageants de culture de lymphocytes B d'un patient

atteint de SEP, immortalisés par la souche B95 du virus d'Epstein-Barr (EBV) et exprimant des particules rétrovirales associées à une activité transcriptase inverse telle que décrite par Perron et coll. (3) et dans  
5 les demandes de brevets français SEP 10, 11 et 12. le clone LB19 dont la séquence identifiée par SEQ ID NO 59, est présentée dans la figure 35.

Ce clone permet de définir, avec le clone GM3 préalablement séquencé et le clone G+E+A (cf Exemple 15),  
10 une région de 690 paires de bases représentative d'une partie significative du gène gag du rétrovirus MSRV-1, telle que présentée dans la figure 36. Cette séquence dénommée SEQ ID NO 88 est reconstituée à partir de différents clones se recouvrant à leurs extrémités. Cette  
15 séquence est identifiée sous la dénomination région MSRV-1 "gag\*". Dans la figure 36, une trame de lecture potentielle avec la traduction en acides aminés est présentée en-dessous de la séquence en acides nucléiques.

20 **EXEMPLE 13 : OBTENTION D'UN CLONE FBd13 CONTENANT UNE REGION DE GENE pol APPARENTEE AU RETROVIRUS MSRV-1 ET UNE REGION ENV APPAREMMENT INCOMPLETE CONTENANT UN CADRE DE LECTURE (ORF) POTENTIEL POUR UNE GLYCOPROTEINE.**

25 Extraction des ARN viraux: les ARN ont été extraits selon la méthode brièvement décrite ci-après.

Un "pool" de surnageant de culture de lymphocytes B de patients atteints de SEP (650ml) est centrifugé 30 minutes à 10 000 g. Le culot viral obtenu  
30 est resuspendu dans 300 microlitres de PBS/10mM MgCl<sub>2</sub>. Le matériel est traité par un mélange DNase (100mg/ml)/RNase (50mg/ml) 30 minutes à 37°C puis par de la protéinase K (50mg/ml) 30 minutes à 46°C.

Les acides nucléiques sont extraits par un volume  
35 d'un mélange phénol/SDS 0,1% (V/V) chauffé à 60°C puis réextraits par un volume de phénol/chloroforme (1/1; V/V).

La précipitation du matériel est effectuée par 2,5 V d'éthanol en présence de 0,1 V d'acétate de sodium pH=5,2. Le culot obtenu après centrifugation est resuspendu dans 50 microlitres d'eau DEPC stérile.

- 5 L'échantillon est de nouveau traité par 50mg/ml de DNase "RNase free" 30 minutes à température ambiante, extrait par un volume de phénol/chloroforme, et précipité en présence d'acétate de sodium et d'éthanol.

- 10 L'ARN obtenu est quantifié par une lecture de D.O. à 260 nm. La présence de MSRV-1 et l'absence de contaminant ADN est contrôlée par une PCR et une RTPCR MSRV-1 spécifique associée à un ELOSA spécifique du génome MSRV-1.

15 Synthèse de cDNA:

- 5 mg d'ARN sont utilisés pour synthétiser un cDNA amorcé par un oligonucléotide polydT selon les instructions du kit "cDNA Synthesis Module" (ref RPN 1256, Amersham) avec quelques modifications: la  
20 rétrotranscription est effectuée à 45°C au lieu des 42°C recommandés.

Le produit de synthèse est purifié par une double extraction et une double purification suivant les instructions du fabricant.

- 25 La présence de MSRV-1 est vérifiée par une PCR MSRV-1 associée à un ELOSA spécifique du génome MSRV-1.

"Long Distance PCR": (LD-PCR)

- 500 ng de cDNA sont utilisés pour l'étape de LD-  
30 PCR (Expand Long Template System; Boehringer (ref.1681 842)).

Plusieurs couples d'oligonucléotides ont été utilisés. Parmi ceux-ci, le couple défini par les amorces suivantes :

- 35 amorce 5': GGAGAAGAGC AGCATAAGTG G (SEQ ID N°66)  
amorce 3': GTGCTGATTG GTGTATTTAC AATCC (SEQ ID N°67).

Les conditions d'amplification sont les suivantes: 94°C 10 secondes

56°C 30 secondes

5 68°C 5 minutes ;

10 cycles puis 20 cycles avec un incrément de 20 secondes à chaque cycle sur le temps d'élongation. A l'issue de cette première amplification, 2 microlitres du produit d'amplification sont soumis à une deuxième  
10 amplification dans les mêmes conditions que précédemment.

Les LD-PCR sont conduites dans un appareil PCR Perkin modèle 9600 dans des microtubes à paroi fine (Boehringer).

15 Les produits d'amplification sont contrôlés par électrophorèse de 1/5 du volume d'amplification (10 microlitres) en gel d'agarose 1%. Pour le couple d'amorces décrit ci-dessus, on obtient une bande d'environ 1,7 Kb.

20 Clonage du fragment amplifié:

Le produit PCR a été purifié par passage sur un gel d'agarose préparatif puis sur une colonne Costar (Spin; D. Dutcher) selon les instructions du fournisseur.

2 microlitres de la solution purifiée sont  
25 raboutés avec 50 ng de vecteur PCRII selon les instructions du fournisseur (TA Cloning Kit; British Biotechnology)).

Le vecteur recombinant obtenu est isolé par transformation de bactéries DH5aF' compétentes. Les  
30 bactéries sont sélectionnées sur leur résistance à l'ampicilline et la perte du métabolisme pour le Xgal (=colonies blanches). La structure moléculaire du vecteur recombinant est confirmée par minipréparation plasmidique et hydrolyse par l'enzyme EcoRI.

35 FBd13, un clone positif pour tous ces critères a été sélectionné. Une préparation à large échelle du

plasmide recombinant a été effectué à l'aide du kit Midiprep Quiagen (ref 12243) selon les instructions du fournisseur.

Le séquençage du clone FBd13 est effectué grâce  
5 au kit Prism Ready Amplitaq FS dye terminator Perkin (ref. 402119) suivant les instructions du fabricant. Les réactions de séquences sont déposées sur un séquenceur automatique Perkin de type 377 ou 373A. La stratégie de séquençage consiste en une "marche sur le gène" réalisée  
10 sur les deux brins du clone FBd13.

La séquence du clone FBd13 est identifiée par SEQ ID NO 58.

Dans la figure 37, l'homologie de séquence entre le clone FBd13 et le rétrovirus HSERV-9 est représentée  
15 sur le tableau matriciel par un trait plein, pour toute homologie partielle supérieure ou égale à 70 %. On peut constater que des homologies existent dans les régions flanquantes du clone (avec le gène *pol* en 5' et avec le gène *env* puis le LTR en 3'), mais que la région interne  
20 est totalement divergente et ne présente aucune homologie, même faible, avec le gène *env* d'HSERV-9. De plus, il apparaît que le clone FBd13 contient une région "env" plus longue que celle qui est décrite pour l'endogène défectif HSERV-9 ; on peut ainsi constater que la région divergente  
25 interne constitue un "insert" entre les régions d'homologie partielle avec les gènes défectifs HSERV-9.

Cette séquence supplémentaire détermine une orf potentielle, dénommée ORF B13, qui est représentée par sa séquence aminoacide SEQ ID NO 87.

30 La structure moléculaire du clone FBd13 a été analysée à l'aide du logiciel GeneWork et des banques de données Genbank et SwissProt.

5 sites de glycosylation ont été trouvés.

La protéine n'a pas d'homologie significative  
35 avec des séquences déjà connues.



Il est probable que ce clone provienne d'une recombinaison avec un élément rétroviral endogène (ERV), liée à la réplication de MSRV-1.

Un tel phénomène n'est pas sans générer l'expression de polypeptides, voire de protéines rétrovirales endogènes qui ne sont pas nécessairement tolérées par le système immunitaire. Un tel schéma d'expression aberrante d'éléments endogènes apparentés à MSRV-1 et/ou induite par ce dernier est susceptible de multiplier les antigènes aberrants et, donc, de contribuer à l'induction de processus autoimmuns tels qu'on les observe dans la SEP. Il constitue à l'évidence un élément original, jamais décrit à ce jour. En effet, l'interrogation des banques de données de séquences nucléiques disponibles dans la version n° 19 (1996) du logiciel "Entrez" (NCBI, NIH, Bethesda, USA) n'a pas permis d'identifier de séquence homologue connue comprenant l'ensemble de la région env de ce clone.

**EXEMPLE 14 : OBTENTION D'UN CLONE FP6 CONTENANT UNE PARTIE DU GENE *pol*, AVEC UNE REGION CODANT POUR L'ENZYME TRANSCRIPTASE INVERSE HOMOLOGUE AU CLONE POL\* MSRV-1, ET UNE REGION 3'*pol* DIVERGENTE DES SEQUENCES EQUIVALENTES DECRITES DANS LES CLONES POL\*, *tpol*, *FBd3*, *JLBc1* et *JLBc2*.**

Une 3'RACE a été effectuée sur de l'ARN total extrait de plasma d'un patient atteint de SEP. Un plasma témoin sain, traité sous les mêmes conditions, a été utilisé comme contrôle négatif. La synthèse de cDNA a été réalisée avec l'amorce oligo dT modifiée suivante:

5' GACTCGCTGC AGATCGATTT TTTT TTTT 3' (SEQ ID NO 68)

et la transcriptase inverse "Expand RT" de Boehringer selon les conditions préconisées par la société. Une PCR a été effectuée avec l'enzyme KlenTag (Clontech) sous les conditions suivantes : 94°C 5 min puis

93°C 1 min, 58°C 1 min, 68°C 3 min pendant 40 cycles et 68°C pendant 8 min et avec un volume réactionnel final de 50 µl.

Amorces utilisées pour la PCR :

- 5           - amorce 5', identifiée par SEQ ID NO 69  
5' GCCATCAAGC CACCCAAGAA CTCTTAACTT 3' ;  
- amorce 3', identifiée par SEQ ID NO 68 (=la même que pour le cDNA)

10           Une deuxième PCR dite "semi-nichée" a été réalisée avec une amorce 5' située à l'intérieur de la région déjà amplifiée. Cette deuxième PCR a été effectuée sous les mêmes conditions expérimentales que celles utilisées lors de la première PCR, en utilisant 10 µl du produit d'amplification issu de la première PCR.

- 15           Amorces utilisées pour la PCR semi-nichée:  
- amorce 5', identifiée par SEQ ID NO 70  
5' CCAATAGCCA GACCATTATA TACACTAATT 3' ;  
- amorce 3', identifiée par SEQ ID NO 68 (=la même que pour le cDNA)

20           Les amorces SEQ ID NO 69 et SEQ ID NO 70 sont spécifiques de la région pol\* : position n°403 à n°422 et n°641 à n°670 respectivement.

25           Un produit d'amplification a ainsi été obtenu à partir de l'ARN extracellulaire extrait du plasma d'un patient atteint de SEP. Le fragment correspondant n'a pas été observé pour le plasma d'un témoin sain. Ce produit d'amplification a été cloné de la façon suivante.

30           L'ADN amplifié a été inséré dans un plasmide à l'aide du kit TA Cloning<sup>TM</sup>. Les 2 µl de solution d'ADN ont été mélangés avec 5 µl d'eau distillée stérile, 1 µl d'un tampon de ligation 10 fois concentré "10X LIGATION BUFFER", 2 µl de "PCR<sup>TM</sup> VECTOR" (25 ng/ml) et 1 µl de "TA DNA LIGASE". Ce mélange a été incubé une nuit à 12°C. Les étapes suivantes ont été réalisées conformément au  
35 instructions du kit TA Cloning® (British Biotechnology). A la fin de la procédure, les colonies blanches de bactéries

recombinantes (white) ont été repiquées pour être cultivées et permettre l'extraction des plasmides incorporés selon la procédure dite de "miniprep" (17). La préparation de plasmide de chaque colonie recombinante a  
5 été coupée par une enzyme de restriction appropriée et analysée sur gel d'agarose. Les plasmides possédant un insert détecté sous lumière UV après marquage du gel au bromure d'éthidium, ont été sélectionnés pour le séquençage de l'insert, après hybridation avec une amorce  
10 complémentaire du promoteur Sp6 présent sur le plasmide de clonage du TA cloning kit®. La réaction préalable au séquençage a ensuite été effectuée selon la méthode préconisée pour l'utilisation du kit de séquençage "Prism ready reaction kit dye deoxyterminator cycle sequencing  
15 kit" (Applied Biosystems, ref. 401384) et le séquençage automatique a été réalisé avec l'appareil "Automatic Sequencer, modèle 373 A" Applied Biosystems, selon les instructions du fabricant.

Le clone obtenu, nommé FP6, permet de définir une  
20 région de 467 pb homologue à 89% à la région pol\* du rétrovirus MSRV-1, et une région de 1167 pb homologue à 64% à la région pol d'ERV-9 (n°1634 à 2856).

Le clone FP6 est représenté sur la figure 38 par sa séquence nucléotidique identifiée par SEQ ID NO 61. Les  
25 trois trames de lecture potentielles de ce clone sont présentées par leur séquence aminoacide sous la séquence nucléotidique.

**EXEMPLE 15 : OBTENTION D'UNE REGION DENOMMEE**  
30 **G+E+A CONTENANT UNE ORF POUR UNE PROTEASE RETROVIRALE PAR AMPLIFICATION PCR DE LA SEQUENCE NUCLEIQUE CONTENUE ENTRE LA REGION 5' DEFINIE PAR LE CLONE "GM3" ET LA REGION 3' DEFINIE PAR LE CLONE POL\*, A PARTIR DE L'ARN EXTRAIT D'UN POOL DE PLASMAS DE PATIENTS ATTEINTS DE SEP.**

Des oligonucléotides spécifiques des séquences MSRV-1 déjà identifiées par la Demanderesse ont été définis pour amplifier l'ARN rétroviral provenant de virions présents dans le plasma de patients atteints de SEP. Des réactions de contrôle ont été effectuées de manière à contrôler la présence de contaminants (réaction sur de l'eau). L'amplification consiste en une étape de RT-PCR, suivie d'une PCR "nichée". Des couples d'amorces ont été définis pour amplifier trois régions chevauchantes (dénommées G, E et A) sur les régions définies par les séquences des clones GM3 et pol\*, préalablement décrits.

RT-PCR semi nichée pour l'amplification de la région G:

-dans le premier cycle RT-PCR, les amorces suivantes sont utilisées :

amorce 1 : SEQ ID NO 71 (sens)

amorce 2 : SEQ ID NO 72 (anti-sens)

-dans le deuxième cycle PCR, les amorces suivantes sont utilisées :

amorce 1 : SEQ ID NO 73 (sens)

amorce 4 : SEQ ID NO 74 (anti-sens)

RT-PCR nichée pour l'amplification de la région E:

-dans le premier cycle RT-PCR, les amorces suivantes sont utilisées :

amorce 5 : SEQ ID NO 75 (sens)

amorce 6 : SEQ ID NO 76 (anti-sens)

-dans le deuxième cycle PCR, les amorces suivantes sont utilisées :

amorce 7 : SEQ ID NO 77 (sens)

amorce 8 : SEQ ID NO 78 (anti-sens)

RT-PCR semi nichée pour l'amplification de la région A:

-dans le premier cycle RT-PCR, les amorces suivantes sont utilisées :

amorce 9 : SEQ ID NO 79 (sens)

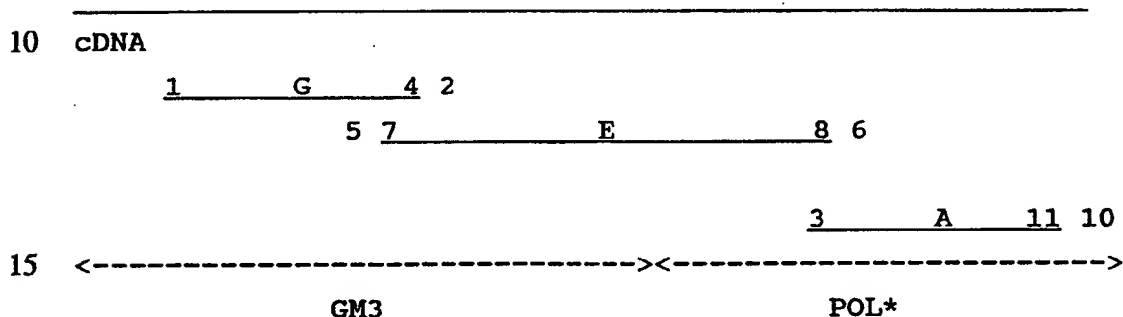
amorce 10 : SEQ ID NO 80 (anti-sens)

-dans le deuxième cycle PCR, les amorces suivantes sont utilisées :

amorce 9 : SEQ ID NO 81(sens)

amorce 11 :SEQ ID NO 82 (anti-sens)

5 Les amorces et les régions G, E et A qu'elles définissent sont positionnées comme suit :



La séquence de la région définie par les différents clones G, E et A a été déterminée après clonage et séquençage des produits d'amplification "nichée".

Les clones G, E et A ont été rassemblés par PCR avec les amorces 1 en 5' du fragment G et 11 en 3' du fragment A, précédemment décrites. Un fragment G+E+A d'environ 1580bp a été amplifié et inséré dans un plasmide à l'aide du kit TA Cloning (marque de commerce). La séquence du produit d'amplification correspondant à G+E+A a été déterminée et l'analyse des recouvrements G+E et E+A a été réalisée. La séquence est représentée dans la figure 39, et correspond à la séquence SEQ ID NO 89.

30 Une trame de lecture codant pour une protéase rétrovirale MSRV-1 a été trouvée dans la région E. La séquence aminoacide de la protéase, identifiée par SEQ ID NO 90, est présentée dans la figure 40.

35 **EXEMPLE 16 : OBTENTION D'UN CLONE LTRGAG12, APPARENTE AU UN ELEMENT RETROVIRAL ENDOGENE (ERV) PROCHE**

DE MSRV-1, DANS L'ADN D'UNE LIGNEE LYMPHOBLASTOIDE DE SEP PRODUISANT DES VIRIONS ET EXPRIMANT LE RETROVIRUS MSRV-1.

Une PCR nichée a été effectuée sur l'ADN extrait  
5 d'une lignée lymphoblastoïde (lymphocytes B immortalisés par le virus EBV, souche B95, comme décrit précédemment et comme cela est bien connu de l'homme de l'art) exprimant le rétrovirus MSRV-1 et provenant des lymphocytes du sang périphérique d'un malade atteint de SEP.

10 Dans la première étape PCR, les amorces suivantes sont utilisées :

amorce 4327 : CTCGATTTCT TGCTGGGCCT TA (SEQ ID NO 83)

amorce 3512 : GTTGATTCCC TCCTCAAGCA (SEQ ID NO 84)

Cette étape comprend 35 cycles d'amplification  
15 avec les conditions suivantes: 1 min à 94°C, 1 min à 54°C et 4 min à 72°C.

Dans le deuxième étape PCR, les amorces suivantes sont utilisées :

amorce 4294 : CTCTACCAAT CAGCATGTGG (SEQ ID NO 85)

20 amorce 3591 : TGTTCCTCTT GGTCCCTAT (SEQ ID NO 86)

Cette étape comprend 35 cycles d'amplification avec les conditions suivantes: 1 min à 94°C, 1 min à 54°C et 4 min à 72°C.

Les produits issus de la PCR ont été purifiés  
25 après purification sur gel d'agarose selon les méthodes conventionnelles (17), puis resuspendus dans 10 ml d'eau distillée. Une des propriétés de la polymérase Taq consistant à ajouter une adénine à l'extrémité 3' de chacun des deux brins d'ADN, l'ADN obtenu a été directement  
30 inséré dans un plamide à l'aide du kit TA Cloning™ (British Biotechnology). Les 2 µl de solution d'ADN ont été mélangés avec 5 µl d'eau distillée stérile, 1 µl d'un tampon de ligation 10 fois concentré "10X LIGATION BUFFER", 2 µl de "PCR™ VECTOR" (25 ng/ml) et 1 µl de  
35 "TA DNA LIGASE". Ce mélange a été incubé une nuit à 12°C. Les étapes suivantes ont été réalisées conformément au

instructions du kit TA Cloning® (British Biotechnology). A la fin de la procédure, les colonies blanches de bactéries recombinantes (white) ont été repiquées pour être cultivées et permettre l'extraction des plasmides incorporés selon la procédure dite de "miniprep" (17). La préparation de plasmide de chaque colonie recombinante a été coupée par une enzyme de restriction appropriée et analysée sur gel d'agarose. Les plasmides possédant un insert détecté sous lumière UV après marquage du gel au bromure d'éthidium, ont été sélectionnés pour le séquençage de l'insert, après hybridation avec une amorce complémentaire du promoteur Sp6 présent sur le plasmide de clonage du TA Cloning Kit®. La réaction préalable au séquençage a ensuite été effectuée selon la méthode préconisée pour l'utilisation du kit de séquençage "Prism ready reaction kit dye deoxyterminator cycle sequencing kit" (Applied Biosystems, ref. 401384) et le séquençage automatique a été réalisé avec l'appareil "Automatic Sequencer, modèle 373 A" Applied Biosystems selon les instructions du fabricant.

Ainsi, un clone dénommé LTRGAG12 a pu être obtenu et est représenté par sa séquence interne identifiée par SEQ ID NO 60.

Ce clone est vraisemblablement représentatif d'éléments endogènes proches d'ERV-9, présent dans l'ADN humain, notamment dans l'ADN de patients atteints de SEP, et pouvant interférer avec l'expression du rétrovirus MSRV-1, donc pouvant avoir un rôle dans la pathogénie associée au rétrovirus MSRV-1 et pouvant servir de marqueur d'une expression spécifique dans la pathologie concernée.

**EXEMPLE 17 : DETECTION D'ANTICORPS SPECIFIQUES ANTI-MSRV-1 DANS LE SERUM HUMAIN.**

L'identification de la séquence du gène *pol* du rétrovirus MSRV-1 et d'un cadre de lecture ouverte de ce gène a permis de déterminer la séquence SEQ ID NO 63 en acides aminés d'une région dudit gène, référencée par  
5 SEQ ID NO 62.

Différents peptides synthétiques correspondant à des fragments de la séquence protéique de la transcriptase inverse MSRV-1 codée par le gène *pol*, ont été testés pour leur spécificité antigénique vis à vis de sera de patients  
10 atteints de SEP et de témoins sains.

Les peptides ont été synthétisés chimiquement par synthèse en phase solide, selon la technique de Merrifield (22). Les modalités pratiques sont celles décrites ci-après.

15 a) Synthèse des peptides:

Les peptides ont été synthétisés sur une résine phénylacétamidométhyle (PAM)/polystyrène/divinylbenzène (Applied Biosystems, Inc. Foster City, CA), en utilisant un synthétiseur automatique "Applied Biosystems 430A". Les  
20 acides aminés sont couplés sous forme d'esters d'hydroxybenzotriazole (HOBt). Les acides aminés utilisés proviennent de Novabiochem (Laüflerlfingen, Suisse) ou de Bachem (Bubendorf, Suisse).

La synthèse chimique a été effectuée en utilisant  
25 un protocole de double couplage avec la N-méthylpyrrolidone (NMP) comme solvant. Les peptides ont été coupés de la résine ainsi que les protections latérales, de manière simultanée, à l'aide d'acide fluorhydrique (HF) dans un appareil approprié (appareil de coupure de type I,  
30 Peptide Institute, Osaka, Japon).

Pour 1 g de peptidylrésine, 10 ml de HF, 1 ml d'anisole et 1 ml de diméthylsulfure 5DMS sont utilisés. Le mélange est agité durant 45 minutes à -2°C. Le HF est alors évaporé sous vide. Après lavages intensifs à  
35 l'éther, le peptide est élué de la résine par de l'acide acétique 10%, puis lyophilisé.



Les peptides sont purifiés par chromatographie liquide haute performance préparative sur une colonne VYDAC de type C18 (250 x 21 mm) (The Separation Group, Hesperia, CA, USA). L'élution est réalisée par un gradient d'acétonitrile à un débit de 22 ml/min. Les fractions collectées sont contrôlées par une élution en condition isocratique sur une colonne VYDAC® C18 analytique (250 x 4,6 mm), à un débit de 1 ml/min. Les fractions qui présentent le même temps de rétention sont réunies et lyophilisées. La fraction majoritaire est ensuite analysée par chromatographie liquide haute performance analytique, avec le système décrit précédemment. Le peptide qui est considéré comme de pureté acceptable se traduit par un pic unique représentant 95 % minimum du chromatogramme.

Les peptides purifiés sont ensuite analysés dans le but de contrôler leur composition en acides aminés, à l'aide d'un analyseur d'acides aminés automatique Applied Biosystems 420H. La mesure de la masse moléculaire chimique (moyenne) des peptides est obtenue en utilisant la spectrométrie de masse L.S.I.M.S. en mode d'ion positif sur un instrument à double focalisation VG. ZAB.ZSEQ relié à un système d'acquisition DEC-VAX 2000 (VG analytical Ltd, Manchester, Angleterre).

La réactivité des différents peptides a été testée contre des séras de patients atteints de SEP et contre des séras de témoins sains. Ceci a permis de sélectionner un peptide dénommé S24Q dont la séquence est identifiée par SEQ ID NO 63, codé par une séquence nucléotidique du gène pol de MSRV-1 (SEQ ID NO 62).

b) Propriétés antigéniques :

Les propriétés antigéniques du peptide S24Q ont été mises en évidence selon le protocole ELISA décrit ci-dessous.

Le peptide S24Q lyophilisé a été dissous à une concentration de 1 mg/ml dans de l'acide acétique à 10%. Cette solution-mère a été aliquotée et gardée à +4°C pour

- usage sous quinzaine, ou congelée à  $-20^{\circ}\text{C}$ , pour un usage dans les 2 mois. Un aliquot est dilué dans une solution de PBS (Phosphate Buffer Saline) afin d'obtenir une concentration finale de peptide de 5 microgrammes/ml.
- 5 100 microlitres de cette dilution sont déposés dans chaque puits de plaques de microtitration Nunc Maxisorb (nom commercial). Les plaques sont recouvertes d'un adhésif de type "plate-sealer" et maintenues 2 heures à  $+37^{\circ}\text{C}$ , pour la phase d'adsorption du peptide sur le plastique.
- 10 L'adhésif est enlevé et les plaques sont lavées trois fois avec un volume de 300 microlitres d'une solution A (PBS 1x, 0,05% Tween 20®), puis retournées sur un tissu absorbant. Les plaques ainsi égouttées sont remplies avec 250 microlitres par puits d'une solution B (solution A +
- 15 10 % de sérum de chèvre), puis recouvertes d'un adhésif et incubées 1 heure, à  $37^{\circ}\text{C}$ . Les plaques sont ensuite lavées trois fois avec la solution A, comme décrit précédemment.

- Les échantillons de sérum à tester sont préalablement dilués au 1/100ème dans la solution B et
- 20 100 microlitres de chaque sérum dilué à tester sont déposés dans les puits de chaque plaque de microtitration. Un contrôle négatif est déposé dans un puits de chaque plaque, sous la forme de 100 microlitres de tampon B. Les plaques recouvertes d'un adhésif sont alors incubées
- 25 1 heure 30 min. à  $37^{\circ}\text{C}$ . Les plaques sont ensuite lavées trois fois avec la solution A, comme décrit précédemment. Pour la réponse IgG, un anticorps de chèvre marqué à la peroxydase et dirigé contre les IgG humaines (commercialisé par Jackson Immuno Research Inc.) est dilué
- 30 dans la solution B (dilution 1/10 000). 100 microlitres de la dilution adéquate de l'anticorps marqué sont alors déposés dans chaque puits des plaques de microtitration et les plaques recouvertes d'un adhésif sont incubées 1 heure, à  $37^{\circ}\text{C}$ . Un nouveau lavage des plaques est ensuite
- 35 effectué comme décrit précédemment. Parallèlement, le substrat de la peroxydase est préparé selon les

indications des kits bioMérieux. 100 microlitres de solution-substrat sont déposés dans chaque puits et les plaques sont placées à l'abri de la lumière pendant 20 à 30 minutes, à température ambiante.

5 Une fois la réaction colorée stabilisée, 50 microlitres de Color 2 (nom commercial-bioMérieux) sont déposés dans chaque puits pour stopper la réaction. Les plaques sont immédiatement placées dans un lecteur spectrophotométrique de plaques ELISA, et la densité  
10 optique (D.O.) de chaque puits est lue à une longueur d'onde de 492 nm.

Les échantillons sérologiques sont déposés en double ou en triple et la densité optique (D.O.) correspondant au sérum testé est calculée en faisant la  
15 moyenne des D.O. obtenues pour un même échantillon, à la même dilution.

La D.O. nette de chaque sérum correspond à la D.O. moyenne du sérum de laquelle est soustraite la D.O. moyenne du témoin négatif (solution B : PBS, Tween 20®  
20 0,05%, 10% sérum de chèvre).

c) Détection d'anticorps IgG anti-MSRV-1 (S24Q) par ELISA :

La technique décrite ci-dessus a été utilisée avec le peptide S24Q pour rechercher la présence  
25 d'anticorps IgG spécifiques anti-MSRV-1 dans le sérum de 15 patients, pour lesquels un diagnostic certain de SEP a été posé selon les critères de Poser (23), et de 15 témoins sains (donneurs de sang).

Dans la figure 41, les résultats de chaque sérum  
30 testé avec un anticorps anti-IgG, sont représentés. Chaque barre verticale représente la densité optique (D.O à 492 nm) nette d'un sérum testé. L'axe des ordonnées indique la D.O. nette au sommet des barres verticales. Les 15 premières barres verticales situées à gauche de la  
35 ligne pointillée verticale, représentent les sérums de 15 témoins sains (donneurs de sang) et les 15 barres

verticales situées à droite de la ligne pointillée verticale, représentent les sérums de 15 cas de SEP testés. Le graphique permet de visualiser 2 témoins dont la D.O. émerge au dessus des valeurs groupées de la population témoin. Ces valeurs peuvent représenter la présence d'IgG spécifiques chez des patients séropositifs non-malades. Deux méthodes ont donc été évaluées pour déterminer le seuil statistique de positivité du test.

La moyenne des D.O. nettes des témoins, y compris les témoins avec des D.O. nettes élevées, est de 0,129 et l'écart type est de 0,06. Sans les 2 témoins dont les D.O. sont supérieures à 0,2, la moyenne des témoins "négatifs" est de 0,107 et l'écart-type de 0,03. Un seuil de positivité théorique peut être calculé selon la formule :

valeur seuil (moyenne des D.O. nettes des témoins négatifs) + (2 ou 3 x écart-type des D.O. nettes des témoins négatifs).

Dans le premier cas, on considère qu'il existe des séropositifs non-malades et la valeur seuil est égale à :  $0,11 + (3 \times 0,03) = 0,20$ . Les résultats négatifs représentent un "bruit de fond" non spécifique de la présence d'anticorps spécifiquement dirigés contre un épitope du peptide.

Dans le deuxième cas, si l'ensemble des témoins constitués de donneurs de sang en bonne santé apparente est pris comme base de référence, sans exclure les sérums a priori séropositifs, l'écart-type des "témoins non-SEP" est de 0,116. La valeur seuil devient alors :  $0,13 + (3 \times 0,06) = 0,31$ .

Selon cette dernière analyse, le test est spécifique de la SEP. A cet égard, on constate que le test est spécifique de la SEP, puisque comme montré dans le tableau 1, aucun témoin n'a une D.O. nette au-dessus de ce seuil. En fait ce résultat reflète le fait que les titres d'anticorps chez les patients atteints de SEP sont en

majorité plus élevés que chez des témoins sains ayant été en contact avec MSRV-1.

En fonction du premier mode de calcul et comme représenté à la figure 41 et dans le tableau 3, 6 des 15 sérums SEP donnent un résultat positif (D.O. supérieure ou égale à 0,2) indiquant la présence d'IgG spécifiquement dirigées contre le peptide S24Q, donc contre une partie de l'enzyme transcriptase inverse du rétrovirus MSRV-1 codée par son gène pol et, par conséquent, contre le rétrovirus MSRV-1.

Ainsi, environ 40 % des patients SEP testés ont réagi contre un épitope porté par le peptide S24Q et présentent des IgG circulantes dirigées contre ce dernier.

2 donneurs de sang, en apparente bonne santé, sur 15 présentent un résultat positif. Ainsi, il apparaît qu'environ 13 % de la population non-malade peut avoir été en contact d'un épitope porté par le peptide S24Q dans des conditions ayant conduit à une immunisation active qui se traduit par la persistance d'IgG sériques spécifiques. Ces conditions sont compatibles avec une immunisation contre la transcriptase inverse du rétrovirus MSRV-1, lors d'une infection par le (et/ou réactivation du) rétrovirus MSRV-1. L'absence de pathologie neurologique apparente évoquant la SEP chez ces témoins séropositifs peut indiquer qu'ils sont porteurs sains, ont éliminé un virus infectieux après s'être immunisés ou qu'ils constituent une population à risque de porteurs chroniques. En effet, les données épidémiologiques montrant qu'un agent pathogène présent dans l'environnement des régions à haute prévalence de SEP, peut être la cause de cette maladie, impliquent qu'une fraction de la population exempte de SEP a nécessairement été en contact avec un tel agent pathogène. On a montré que le rétrovirus MSRV-1 constitue tout ou partie de cet "agent pathogène" à l'origine de la SEP et il est donc normal que des témoins pris dans une

population saine présentent des anticorps de type IgG contre des composants du rétrovirus MSRV-1.

Enfin, la détection d'anticorps anti-S24Q chez seulement une SEP sur deux testée ici, peut refléter le fait que ce peptide ne représente pas un épitope immunodominant MSRV-1, que des variations inter-individuelles de souches peuvent induire une immunisation contre un motif peptidique divergent dans la même région, ou encore que l'évolution de la maladie et les traitements suivis peuvent moduler dans le temps la réponse anticorps contre le peptide S24Q.

TABLEAU N°3

TEMOINS	SEP
0,101	0,136
0,058	0,391
0,126	0,37
0,131	0,119
0,105	0,267
0,294	0,141
0,116	0,102
0,088	0,18
0,105	0,411
0,172	0,164
0,137	0,049
0,223	0,644
0,08	0,268
0,073	0,065
0,132	0,074
Moyenne 0,129	
Ecart Type 0,06	
Seuil 0,31	

15

d) Détection d'anticorps IgM anti-MSRV-1 par ELISA :

La technique ELISA avec le peptide S24Q a été utilisée pour rechercher la présence d'anticorps spécifiques IgM anti-MSRV-1 dans les mêmes sérums que précédemment.

20

Dans la figure 42, les résultats de chaque sérum testé avec un anticorps anti-IgM, sont représentés. Chaque barre verticale représente la densité optique (D.O à 492 nm) nette d'un sérum testé. L'axe des ordonnées  
5 indique la D.O. nette au sommet des barres verticales. Les 15 premières barres verticales situées à gauche de la ligne verticale coupant l'axe des abscisses représentent les sérums de 15 témoins sains (donneurs de sang) et les barres verticales situées à droite de la ligne pointillée  
10 verticale, représentent les sérums de 15 cas de SEP testés.

La moyenne des D.O. des SEP testées est de : 1,6.

La moyenne des D.O. nettes des témoins est de : 0,7.

15 L'écart type des témoins négatifs est de : 0,6.

Le seuil de positivité théorique peut être calculé selon la formule :

valeur seuil = (moyenne des D.O. des témoins négatifs) + (3 x écart-type des D.O. des témoins négatifs)

20 La valeur seuil est donc égale à  $0,7 + (3 \times 0,6) = 2,5$  ;

Les résultats négatifs représentent un "bruit de fond" non spécifique de la présence d'anticorps spécifiquement dirigés contre un épitope du peptide.

Selon cette analyse et comme montré à la  
25 figure 42 et dans le tableau 4 correspondant, le test IgM est spécifique de la SEP, puisqu'aucun témoin n'a une D.O. nette au-dessus du seuil. 6 des 15 sérums SEP produisent un résultat IgM positif

La différence de séroprévalence entre les SEP et  
30 la population témoin est extrêmement significative : test "chi-2",  $p < 0,002$ .

Ces résultats sont en faveur d'un rôle étiopathogénique de MSRV-1 dans la SEP.

Ainsi, la détection des anticorps IgM et IgG  
35 contre le peptide S24Q permet d'évaluer, seul ou en combinaison avec d'autres peptides MSRV-1, l'évolution

d'une infection par MSRV-1 et/ou de la réactivation virale de MSRV-1.

TABLEAU N°4

5

TEMOINS	SEP
1,449	0,974
0,371	6,117
0,448	2,883
0,456	1,945
0,885	1,787
2,235	0,273
0,301	1,766
0,138	0,668
0,16	2,603
1,073	0,802
1,366	0,245
0,283	0,147
0,262	2,441
0,585	0,287
0,356	0,589
Moyenne 0,7	
Ecart Type 0,6	
Valeur seuil 2,5	

Il est possible, grâce aux nouvelles découvertes effectuées et aux nouvelles méthodes mises au point par les inventeurs, de permettre la réalisation perfectionnée de tests diagnostiques de l'infection et/ou de la réactivation MSRV-1, et d'évaluer une thérapie dans la SEP et/ou la PR, sur la base de son efficacité à "négativer" la détection de ces agents dans les fluides biologiques des patients. De plus, la détection précoce chez des personnes ne présentant pas encore de signes neurologiques de SEP, ou de signes rhumatologiques de PR, pourrait permettre d'instaurer un traitement d'autant plus efficace sur l'évolution clinique ultérieure qu'il précéderait le stade lésionnel qui correspond à l'apparition des troubles cliniques. Or, à ce jour, un diagnostic de SEP ou de PR ne peut être établi avant l'installation d'une symptomatologie lésionnelle et, donc, aucun traitement



n'est instauré avant l'émergence d'une clinique évocatrice de lésions ddéjà conséquentes. Le diagnostic d'une infection et/ou réactivation de MSRV-1 et/ou MSRV-2 chez l'homme est donc déterminant et la présente invention en  
5 fournit les moyens.

Il est ainsi possible, outre de réaliser un diagnostic de l'infection et/ou de la réactivation MSRV-1, d'évaluer une thérapie dans la SEP sur la base de son efficacité à "négativer" la détection de ces agents dans  
10 les fluides biologiques des patients.

## BIBLIOGRAPHIE

- (1) Norrby E., Prog. Med. Virol., 1978; 24, 1-39.
- (2) Johnson R.T., "Handbook of clinical neurology, 47  
Demyelinating diseases", Vinken P. et Bruyn G.W., eds.  
5 Amsterdam, Elsevier Science Publishing, 1985, 319-336.
- (3) Perron H. et coll., Res. Virol. 1989, 140, 551-561.
- (4) Perron H. et coll., "Current concepts in multiple  
sclerosis" Wiethölter et coll., eds. Amsterdam,  
Elsevier, 1991, 111-116.
- 10 (5) Perron H. et coll., The Lancet 1991, 337, 862-863.
- (6) Perron H. et coll., J. Gen. Virol. 1993, 74, 65-72.
- (7) Fields et Knipe, Fundamental Virology 1986, Rev Press  
N.Y.
- (8) Nielsen P.E. et coll., Science 1991; 254, 1497-1500.
- 15 (9) Maniatis et al, Molecular Cloning, Cold Spring Harbor,  
1982.
- (10) Southern. E.M., J. Mol. Biol. 1975, 98, 503.
- (11) Dunn A.R. et Hassel J.A., Cell 1977, 12, 23.
- (12) Shih et coll., J. Virol. 1989, 63, 64-75.
- 20 (13) Perron H. et coll., Res. Vir. 1992, 143, 337-350.
- (14) Meyerhans et coll., Cell 1989, 58, 901-910.
- (15) Linial M.L. and Miller A.D., "Current topics in  
microbiology and immunobiology. Retroviruses,  
strategies of replication" vol. 157, 125-152;  
25 Swanstrom R. et Vogt P.K., éditeurs, Springer-Verlag,  
Heidelberg 1990.
- (16) Lori F. et coll., J. Virol. 1992, 66, 5067-5074.
- (17) Sambrook J., Fritsch E.F., et Maniatis T., Molecular  
cloning, a laboratory manual. Cold Spring Harbor  
Laboratory Press, 1989.
- 30 (18) La Mantia et coll., Nucleic Acids Research 1991, 19,  
1513-1520.
- (19) Gonzalez-Quintial R, Baccala R, Pope R M and  
Theofilopoulos N, J. Clin. Invest, vol. 97, Number 5,  
35 pp1335-1343, 1996.

- (20) Chomzynski P. et N. Sacchi, Analytical Biochemistry 1987, 162, 156-159.
- (21) F. Mallet et coll., Journal of Clinical Microbiology 1993; 31, 1444-1449.
- 5 (22) G. Barany and R.B. Merrifield, 1980, In the Peptides, 2, 1-284, Gross E and Meienhofer J, Eds Academic Press, New York.
- (23) Poser et al, Gbers G.C. eds. The diagnosis of multiple sclerosis Thieme Stratton Inc, New York 1984 :  
10 225-229.
- (24) La Mantia et coll., Nucleic Acid Research 1989, 17, 5913-22.
- (25) PLAZA, A ; KONO, D.H ; THEOFILOPOULOS, A.N. NEW HUMAN  
V $\beta$  GENES AND POLYMORPHIC VARIANTS. J. Imm; 147(12):  
15 4360-4365, 1991.

## LISTE DE SEQUENCES

## (1) INFORMATIONS GENERALES:

5

## (i) DEPOSANT:

- (A) NOM: BIOMERIEUX  
 (B) RUE: AUCUNE  
 (C) VILLE: MARCY L'ETOILE  
 (E) PAYS: FRANCE  
 (F) CODE POSTAL: 69280

10

## (ii) TITRE DE L'INVENTION:

15

## (iii) NOMBRE DE SEQUENCES: 92

## (iv) FORME DECHIFFRABLE PAR ORDINATEUR:

- (A) TYPE DE SUPPORT: Floppy disk  
 (B) ORDINATEUR: IBM PC compatible  
 (C) SYSTEME D' EXPLOITATION: PC-DOS/MS-DOS  
 (D) LOGICIEL: PatentIn Release #1.0, Version #1.30 (OEB)

20

## 25 (2) INFORMATIONS POUR LA SEQ ID NO: 1:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 1158 paires de bases  
 (B) TYPE: nucléotide  
 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

30

## (ii) TYPE DE MOLECULE: ADNc

35

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 1:

CCCTTTGCCA CTACATCAAT TTTAGGAGTA AGGAAACCCA ACGGACAGTG GAGGTTAGTG 60  
 CAAGAACTCA GGATTATCAA TGAGGCTGTT GTTCCTCTAT ACCCAGCTGT ACCTAACCCT 120  
 TATACAGTGC TTTCCCAAAT ACCAGAGGAA GCAGAGTGGT TTACAGTCCT GGACCTTAAG 180  
 40 GATGCCCTTT TCTGCATCCC TGTACGTCCT GACTCTCAAT TCTTGTTTGC CTTTGAAGAT 240  
 CCTTTGAACC CAACGTCTCA ACTCACCTGG ACTGTTTTAC CCCAAGGGTT CAGGGATAGC 300  
 CCCCATCTAT TTGGCCAGGC ATTAGCCCAA GACTTGAGTC AATTCTCATA CCTGGACACT 360  
 CTTGTCCTTC AGTACATGGA TGATTACTT TTAGTCGCCC GTTCAGAAAC CTTGTGCCAT 420  
 CAAGCCACCC AAGAACTCTT AACTTTCCTC ACTACCTGTG GCTACAAGGT TTCCAAACCA 480  
 45 AAGGCTCGGC TCTGCTCACA GGAGATTAGA TACTNAGGGC TAAATATATC CAAAGGCACC 540  
 AGGGCCCTCA GTGAGGAACG TATCCAGCCT ATACTGGCTT ATCCTCATCC CAAAACCCTA 600  
 AAGCAACTAA GAGGGTTCCT TGGCATAACA GGTTCCTGCC GAAACAGAT TCCCAGGTAC 660  
 ASCCCAATAG CCAGACCATT ATATACACTA ATTANGGAAA CTCAGAAAGC CAATACCTAT 720  
 TTAGTAAGAT GGACACCTAC AGAAGTGGCT TTCCAGGCCC TAAAGAAGGC CCTAACCCAA 780  
 50 GCCCCAGTGT TCAGCTTGCC AACAGGGCAA GATTTTCTT TATATGCCAC AGAAAAACA 840  
 GGAATAGCTC TAGGAGTCCT TACGCAGGTC TCAGGGATGA GCTTGCAACC CGTGGTATAC 900  
 CTGAGTAAGG AAATTGATGT AGTGCCAAAG GGTGGCCTC ATNGTTTATG GGTAATGGNG 960  
 GCAGTAGCAG TCTNAGTATC TGAAGCAGTT AAAATAATAC AGGGAAGAGA TCTTNCTGTG 1020  
 TGGACATCTC ATGATGTGAA CGGCATACTC ACTGCTAAAG GAGACTTGTG GTTGTCAGAC 1080

AACCATTTAC TTAANTATCA GGCTCTATTA CTTGAAGAGC CAGTGCTGNG ACTGCGCACT1140  
TGTGCAACTC TTAAACCC 1158

5 (2) INFORMATIONS POUR LA SEQ ID NO: 2:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 10 (A) LONGUEUR: 297 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

15 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 2:

20 CCCTTTGCCA CTACATCAAT TTAGGAGTA AGGAAACCCA ACGGACAGTG GAGGTTAGTG 60  
CAAGAACTCA GGATTATCAA TGAGGCTGTT GTTCCTCTAT ACCCAGCTGT ACCTAACCCT 120  
TATACAGTGC TTTCCCAAAT ACCAGAGGAA GCAGAGTGGT TTACAGTCCT GGACCTTAAG 180  
GATGCCTTTT TCTGCATCCC TGTACGTCCT GACTCTCAAT TCTTGTTTGC CTTGAAGAT 240  
CCTTTGAACC CAACGTCTCA ACTCACCTGG ACTGTTTTAC CCCAAGGGTT CAAGGGA 297

25 (2) INFORMATIONS POUR LA SEQ ID NO: 3:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 30 (A) LONGUEUR: 85 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 3:

35 GTTTAGGGAT ANCCCTCATC TCTTTGGTCA GGTACTGGCC CAAGATCTAG GCCACTTCTC 60  
AGGTCCAGSN ACTCTGTYCC TTCAG 85

40 (2) INFORMATIONS POUR LA SEQ ID NO: 4:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 45 (A) LONGUEUR: 86 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

50 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 4:

GTTCAGGGAT AGCCCCATC TATTTGGCCA GGCAGTAGCT CAATACTTGA GCCAGTTCTC 60  
ATACCTGGAC AYTCTYGTCC TTCGGT 86

95

## (2) INFORMATIONS POUR LA SEQ ID NO: 5:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

5

- (A) LONGUEUR: 85 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

## (ii) TYPE DE MOLECULE: ADNc

10

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 5:

15 GTTCARRGA TAGCCCCATC TATTTGGCCW RGYATTAGCC CAAGACTTGA GYCAATTCTC 60  
ATACCTGGA CACTTTGTCC TTYRG 85

## (2) INFORMATIONS POUR LA SEQ ID NO: 6:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

20

- (A) LONGUEUR: 85 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

25

## (ii) TYPE DE MOLECULE: ADNc

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 6:

30 GTTCAGGGAT AGCTCCCATC TATTTGGCCT GGCATTAACC CGAGACTTAA GCCAGTTCTY 60  
ATACGTGGAC ACTTTGTCC TTTGG 85

## (2) INFORMATIONS POUR LA SEQ ID NO: 7:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

35

- (A) LONGUEUR: 111 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

40

## (ii) TYPE DE MOLECULE: ADNc

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 7:

45 GTGTTGCCAC AGGGGTTTAR RGATANCYCY CATCTMTTGG GYCWRGYAYT RRCYCRKAY 60  
YTRRGYCAVT TCTYAKRYSY RGSNAYCTB KYCCTTYRGT ACATGGATGA C 111

## (2) INFORMATIONS POUR LA SEQ ID NO: 8:

50

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 645 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple

96

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

5 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 8:

TCAGGGATAG CCCCCATCTA TTTGGCCAGG CATTAGCCCA AGACTTGAGT CAATTCTCAT 60  
 ACCTGGACAC TCTTGTCTT CAGTACATGG ATGATTFACT TTTAGTCGCC CGTTCAGAAA 120  
 CCTTGTGCCA TCAAGCCACC CAAGAACTCT TAACTTTCCT CACTACCTGT GGCTACAAGG 180  
 10 TTTCCAAACC AAAGGCTCGG CTCTGCTCAC AGGAGATTAG ATACTNAGGG CTAAAATTAT 240  
 CCAAAGGCAC CAGGGCCCTC AGTGAGGAAC GTATCCAGCC TATACTGGCT TATCCTCATC 300  
 CCAAACCCT AAAGCAACTA AGAGGGTTCC TTGGCATAAC AGGTTTCTGC CGAAAACAGA 360  
 TTCCAGGTA CASCCTAATA GCCAGACCAT TATATACACT AATTANGGAA ACTCAGAAAG 420  
 CCAATACCTA TTTAGTAAGA TGGACACCTA CAGAAAGTGGC TTTCCAGGCC CTAAAGAAGG 480  
 15 CCCTAACCCA AGCCCCAGTG TTCAGCTTGC CAACAGGGCA AGATTTTTCT TTATATGCCA 540  
 CAGAAAAAAC AGGAATAGCT CTAGGAGTCC TTACGCAGGT CTCAGGGATG AGCTTGCAAC 600  
 CCGTGGTATA CCTGAGTAAG GAAATTGATG TAGTGGCAAA GGGTT 645

20 (2) INFORMATIONS POUR LA SEQ ID NO: 9:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 741 paires de bases  
 (B) TYPE: nucléotide  
 25 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

30 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 9:

CAAGCCACCC AAGAACTCTT AAATTTCTC ACTACCTGTG GCTACAAGGT TTCCAAACCA 60  
 AAGGCTCAGC TCTGCTCACA GGAGATTAGA TACTTAGGGT TAAAATTATC CAAAGGCACC 120  
 AGGGGCCTCA GTGAGGAACG TATCCAGCCT ATACTGGGTT ATCCTCATCC CAAACCCCTA 180  
 35 AAGCAACTAA GAGGGTTCCT TAGCATGATC AGGTTTCTGC CGAAAACAAG ATTCCCAGGT 240  
 ACAACCAAAA TAGCCAGACC ATTATATACA CTAATTAAGG AAATCAGAA AGCCAATACC 300  
 TATTTAGTAA GATGGACACC TAAACAGAAG GCTTTCAGG CCTAAAGAA GGCCCTAACC 360  
 CAAGCCCCAG TGTTAGCTT GCCAACAGGG CAAGATTTT CTTTATATGG CACAGAAAAA 420  
 ACAGGAATCG CTCTAGGAGT CCTTACACAG GTCCGAGGGA TGAGCTTGCA ACCCGTGGCA 480  
 40 TACCTGAATA AGGAAATTGA TGTAGTGGCA AAGGGTTGGC CTCATNGTTT ATGGGTAATG 540  
 GNGGCAGTAG CAGTCTNAGT ATCTGAAGCA GTTAAATATA TACAGGGAAG AGATCTTNCT 600  
 GTGTGGACAT CTCATGATGT GAACGGCATA CTCACTGCTA AAGGAGACTT GTGGTTGTCA 660  
 GACAACCATT TACTTAANTA TCAGGCTCTA TTAAGTGAAG AGCCAGTGCT GNGACTGCGC 720  
 ACTTGTGCAA CTCTTAAACC C 741

45

(2) INFORMATIONS POUR LA SEQ ID NO: 10:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 50 (A) LONGUEUR: 93 paires de bases  
 (B) TYPE: nucléotide  
 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

97

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 10:

5 TGGAAAGTGT TGCCACAGGG CGCTGAAGCC TATCGCGTGC AGTTGCCCGGA TGCCGCCTAT 60  
AGCCTCTACA TGGATGACAT CCTGCTGGCC TCC 93

10 (2) INFORMATIONS POUR LA SEQ ID NO: 11:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 15 (A) LONGUEUR: 96 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

20 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 11:

TTGGATCCAG TGYTGCCACA GGGCGCTGAA GCCTATCGCG TGCAGTTGCC GGATGCCGCC 60  
TATAGCCTCT ACGTGGATGA CCTSCTGAAG CTTGAG 96

25

(2) INFORMATIONS POUR LA SEQ ID NO: 12:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 30 (A) LONGUEUR: 748 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

35

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 12:

40 TGCAAGCTTC ACCGCTTGCT GGATGTAGGC CTCAGTACCG GNGTGCCCCG CGCGCTGTAG 60  
TTGATGTAG AAAGCGCCCG GAAACACGCG GGACCAATGC GTCGCCAGCT TGCGCGCCAG 120  
CGCCTCGTTG CCATTGGCCA GCGCCACGCC GATATCACCC GCCATGGCGC CGGAGAGCGC 180  
CAGCAGACCG GCGGCCAGCG GCGCATTCTC AACGCCGGGC TCGTCGAACC ATTCGGGGGC 240  
GATTTCCGCA CGACCGCGAT GCTGGTTGGA GAGCCAGGCC CTGGCCAGCA ACTGGCACAG 300  
GTTCAGGTAA CCCTGCTTGT CCCGCACCAA CAGCAGCAGG CGGGTCGGCT TGTCGCGCTC 360  
GTCGTGATTG GTGATCCACA CGTCAGCCCC GACGATGGGC TTCACGCCCT TGCCACGCGC 420  
45 TTCCTGTAG ANGCGACCA GCCCGAAGGC ATTGGCGAGA TCGGTCAGCG CCAAGGCGCC 480  
CATGCCATCT TTGGCGGCAG CCTTGAAGGC ATCGTCGAGA CGGACATTGC CATCGACGAC 540  
GGAATATTTC GAGTGGAGAC GGAGGTGGAC GAAGCGCGGC GAATTCATCC GCGTATTGTA 600  
ACGGGTGACA CCTTCGCAA AGCATTCCGG ACGTGCCCGA TTGACCCGGA GCAACCCCGC 660  
ACGGCTGCGC GGGCAGTTAT AATTTCGGCT TACGAATCAA CGGGTTACCC CAGGGCGCTG 720  
50 AAGCCTATCG CGTGCAGTTG CCGGATGC 748



**(2) INFORMATIONS POUR LA SEQ ID NO: 13:**

- 5       (i) CARACTERISTIQUES DE LA SEQUENCE:  
          (A) LONGUEUR: 18 paires de bases  
          (B) TYPE: nucléotide  
          (C) NOMBRE DE BRINS: simple  
          (D) CONFIGURATION: linéaire

- 10       (ii) TYPE DE MOLECULE: ADNc

**(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 13:**

15       GCATCCGGCA ACTGCACG   18

**(2) INFORMATIONS POUR LA SEQ ID NO: 14:**

- 20       (i) CARACTERISTIQUES DE LA SEQUENCE:  
          (A) LONGUEUR: 20 paires de bases  
          (B) TYPE: nucléotide  
          (C) NOMBRE DE BRINS: simple  
          (D) CONFIGURATION: linéaire

- 25       (ii) TYPE DE MOLECULE: ADNc

**(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 14:**

30       GTAGTTCGAT GTAGAAAGCG   20

**(2) INFORMATIONS POUR LA SEQ ID NO: 15:**

- 35       (i) CARACTERISTIQUES DE LA SEQUENCE:  
          (A) LONGUEUR: 18 paires de bases  
          (B) TYPE: nucléotide  
          (C) NOMBRE DE BRINS: simple  
          (D) CONFIGURATION: linéaire

- 40       (ii) TYPE DE MOLECULE: ADNc

**(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 15:**

45       GCATCCGGCA ACTGCACG   18

**(2) INFORMATIONS POUR LA SEQ ID NO: 16:**

- 50       (i) CARACTERISTIQUES DE LA SEQUENCE:  
          (A) LONGUEUR: 23 paires de bases  
          (B) TYPE: nucléotide  
          (C) NOMBRE DE BRINS: simple  
          (D) CONFIGURATION: linéaire

99

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 16:

5 AGGAGTAAG GAAACCCAACG GAC 23

(2) INFORMATIONS POUR LA SEQ ID NO: 17:

10 (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 19 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

15 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 17:

20 TAAGAGTTGC ACAAGTGCG 19

(2) INFORMATIONS POUR LA SEQ ID NO: 18:

25 (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 21 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

30 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 18:

35 TCAGGGATAG CCCCCATCTA T 21

(2) INFORMATIONS POUR LA SEQ ID NO: 19:

40 (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 24 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

45 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 19:

50 AACCCCTTGC CACTACATCA ATTT 24

100

## (2) INFORMATIONS POUR LA SEQ ID NO: 20:

- 5 (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 15 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire
- 10 (ii) TYPE DE MOLECULE: ADNc
- (ix) CARACTERISTIQUES:  
(B) EMPLACEMENT: 5, 7, 10, 13  
(D) AUTRES INFORMATIONS: G représente l'inosine (i)
- 15 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 20:
- GGTCGTGCCG CAGGG 15

20

## (2) INFORMATIONS POUR LA SEQ ID NO: 21:

- 25 (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 21 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc
- 30 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 21:

TTAGGGATA GCCCTCATCTC T 21

35

## (2) INFORMATIONS POUR LA SEQ ID NO: 22:

- 40 (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 21 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc
- 45 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 22:

TCAGGGATAG CCCCCATCTA T 21

50

## (2) INFORMATIONS POUR LA SEQ ID NO: 23:

- (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 24 paires de bases

101

- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

5 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 23

AACCCTTTGC CACTACATCA ATTT 24

10

(2) INFORMATIONS POUR LA SEQ ID NO: 24:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 15 (A) LONGUEUR: 23 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

20 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 24

GCGTAAGGAC TCCTAGAGCT ATT 23

25

(2) INFORMATIONS POUR LA SEQ ID NO: 25:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 30 (A) LONGUEUR: 18 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

35 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 25

TCATCCATGT ACCGAAGG 18

40

(2) INFORMATIONS POUR LA SEQ ID NO: 26:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- 45 (A) LONGUEUR: 20 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

50

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 26

ATGGGGTTCC CAAGTTCCT 20

102

**(2) INFORMATIONS POUR LA SEQ ID NO: 27:****(i) CARACTERISTIQUES DE LA SEQUENCE:**

- 5 (A) LONGUEUR: 20 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

**(ii) TYPE DE MOLECULE: ADNc**

10

**(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 27**

GCCGATATCA CCCGCCATGG 20

15

**(2) INFORMATIONS POUR LA SEQ ID NO: 28:****(i) CARACTERISTIQUES DE LA SEQUENCE:**

- 20 (A) LONGUEUR: 18 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

**(ii) TYPE DE MOLECULE: ADNc**

25

**(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 28**

GCATCCGGCA ACTGCACG 18

30

**(2) INFORMATIONS POUR LA SEQ ID NO: 29:****(i) CARACTERISTIQUES DE LA SEQUENCE:**

- 35 (A) LONGUEUR: 20 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

**(ii) TYPE DE MOLECULE: ADNc**

40

**(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 29**

CGCGATGCTG GTTGGAGAGC 20

45

**(2) INFORMATIONS POUR LA SEQ ID NO: 30:****(i) CARACTERISTIQUES DE LA SEQUENCE:**

- 50 (A) LONGUEUR: 20 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

**(ii) TYPE DE MOLECULE: ADNc**

103

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 30

TCTCCACTCC GAATATTCCG 20

5

(2) INFORMATIONS POUR LA SEQ ID NO: 31:

(i) CARACTERISTIQUES DE LA SEQUENCE:

10

(A) LONGUEUR: 26 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

15

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 31

GATCTAGGCC ACTTCTCAGG TCCAGS 26

20

(2) INFORMATIONS POUR LA SEQ ID NO: 32:

(i) CARACTERISTIQUES DE LA SEQUENCE:

25

(A) LONGUEUR: 23 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

30

(ii) TYPE DE MOLECULE: ADNc

(ix) CARACTERISTIQUES:

(B) EMPLACEMENT: 6, 12, 19

(D) AUTRES INFORMATIONS: G représente l'inosine (i)

35

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 32

CATCTGTTTG GGCAGGCAGT AGC 23

40

2) INFORMATIONS POUR LA SEQ ID NO: 33:

(i) CARACTERISTIQUES DE LA SEQUENCE:

45

(A) LONGUEUR: 24 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

50

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 33

CTTGAGCCAG TTCTCATACC TGGA 24

104

## 2) INFORMATIONS POUR LA SEQ ID NO: 34:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

5

- (A) LONGUEUR: 22 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

## (ii) TYPE DE MOLECULE: ADNc

10

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 34

AGTGYTRCCM CARGGCGCTG AA 22

15

## 2) INFORMATIONS POUR LA SEQ ID NO: 35:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

20

- (A) LONGUEUR: 22 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

## (ii) TYPE DE MOLECULE: ADNc

25

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 35

GMGGCCAGCA GSAKGTATC CA 22

30

## 2) INFORMATIONS POUR LA SEQ ID NO: 36:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

35

- (A) LONGUEUR: 22 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

## (ii) TYPE DE MOLECULE: ADNc

40

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 36

GGATGCCGCC TATAGCCTCT AC 22

45

## 2) INFORMATIONS POUR LA SEQ ID NO: 37:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

50

- (A) LONGUEUR: 22 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

## (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 37

AAGCCTATCG CGTGCA GTTG CC 22

(2) INFORMATIONS POUR LA SEQ ID NO: 38:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 40 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 38:

TAAAGATCTA GAATTCGGCT ATAGGCGGCA TCCGGCAAGT 40

(2) INFORMATIONS POUR LA SEQ ID NO: 39:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 39:

cf figure 28

(2) INFORMATIONS POUR LA SEQ ID NO: 40:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 40:

cf figure 28

(2) INFORMATIONS POUR LA SEQ ID NO: 41:

(i) CARACTERISTIQUES DE LA SEQUENCE:



106

- (A) LONGUEUR: paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

5

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 41:

10 cf figure 34

(2) INFORMATIONS POUR LA SEQ ID NO: 42:

15

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

20

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 42:

25 cf figure 34

(2) INFORMATIONS POUR LA SEQ ID NO: 43:

30

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

35

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 43:

40 cf figure 34

(2) INFORMATIONS POUR LA SEQ ID NO: 44:

45

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

50

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 44:

107

cf figure 34

## (2) INFORMATIONS POUR LA SEQ ID NO: 45:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 45:

## (2) INFORMATIONS POUR LA SEQ ID NO: 46:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 46:

cf figure 13

## (2) INFORMATIONS POUR LA SEQ ID NO: 47:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 23 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 47:

TGATGTGAAC GGCATACTCA CTG 23

## (2) INFORMATIONS POUR LA SEQ ID NO: 48:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 24 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

108

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 48:

5 CCCAGAGGTT AGGAACTCCC TTTC 24

(2) INFORMATIONS POUR LA SEQ ID NO: 49:

10

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 25 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

15

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 49:

20

GCTAAAGGAG ACTTGTGGTT GTCAG 25

(2) INFORMATIONS POUR LA SEQ ID NO: 50:

25

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 22 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

30

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 50:

35

CAACATGGGC ATTCGGATT AG 22

(2) INFORMATIONS POUR LA SEQ ID NO: 51:

40

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

45

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 51:

50

cf figure 15

(2) INFORMATIONS POUR LA SEQ ID NO: 52:

- (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc
- (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 52:  
cf figure 16
- (2) INFORMATIONS POUR LA SEQ ID NO: 53:
- (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc
- (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 53:  
cf figure 17
- (2) INFORMATIONS POUR LA SEQ ID NO: 54:
- (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 21 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc
- (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 54:  
CCTGAGTTCT TGCACTAACC C 21
- (2) INFORMATIONS POUR LA SEQ ID NO: 55:
- (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 23 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc
- (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 55:  
GTCCGTGGG TTTCCTTACT CCT 23

110

## (2) INFORMATIONS POUR LA SEQ ID NO: 56:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- 5 (A) LONGUEUR: paires de bases  
 (B) TYPE: nucléotide  
 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

## 10 (ii) TYPE DE MOLECULE: ADNc

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 56:

cf figure 23

15

## (2) INFORMATIONS POUR LA SEQ ID NO: 57:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- 20 (A) LONGUEUR: paires de bases  
 (B) TYPE: nucléotide  
 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

## 25 (ii) TYPE DE MOLECULE: ADNc

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 57:

cf figure 27 (27a à 27c)

30

## (2) INFORMATIONS POUR LA SEQ ID NO: 58:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- 35 (A) LONGUEUR: 1722 paires de bases  
 (B) TYPE: nucléotide  
 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

## 40 (ii) TYPE DE MOLECULE: ADNc

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 58:

	TGGAGAATAG	CAGCATAAGT	TGGCTGGCAG	AAGTAGGGAA	AGACAGCAAG	50
45	AAGTAAAGAA	AAAAARGAGA	AAGTCAGAGA	AAGAAAAAAA	GAGAGGAAGA	100
	AACAAAGAAG	AACTTGAAGA	GAGAAAGAAG	TAGTAAAGAA	AAAACAGTAT	150
	ACCCTATTCC	TTTAAAAGCC	AGGGTAAATT	TCTGTCTACC	TAGCCAAGGC	200
	ATATTCTTCT	TATGTGGAAC	ATCAACCTAT	ATCTGCCTCC	CCACTAACTG	250
	GACAGGCACC	TGAACCTTAG	TCTTCTAAG	TCCCAACATT	AACATTGCCC	300
50	CAGGAATCA	GACCTATTG	GTACCTGTCA	AAGCTAAAGT	CCCGTCAGTG	350
	CAGAGCCATA	CAACTAATAT	CCCTATTTAT	AGGGTTAGGA	ATGGCTACTG	400
	CTACAGGAAC	TGGAATAGCC	GGTTTATCTA	CTTCATTATC	CTACTACCAT	450
	ACACTCTCAA	AGAATTTCTC	AGACAGTTTG	CAAGAAATAA	TGAAATCTAT	500
	TCTTACTTTA	CAATCCCAAT	TAGACTCTTT	GGCAGCAATG	ACTCTCCAAA	550
55	ACCGCCGAGG	CCCACACCTC	CTCACTGCTG	AGAAAGGAGG	ACTCTGCACC	600
	TTCTTAGGGG	AAGAGTGTTG	TTTTTACACT	AACCAGTCAG	GGATAGTACG	650

111

	AGATGCCACC	TGGCATTAC	AGGAAAGGGC	TTCTGATATC	AGACAATGCC	700
	TTTCAAATC	TTATACCAAC	CTCTGGAGTT	GGGCAACATG	GCTTCTTCCA	750
	TTTCTAGGTC	CCATGGCAGC	CATCTTGCTG	TFACTCACCT	TTGGGCCCTG	800
	TATTTTAAAG	CTTCTTGTC	AATTTGTTTC	CTCTAGGATC	GAAGCCATCA	850
5	AGCTACAGAT	GGTCTTACAA	ATGGAACCCC	AAATGAGTTC	AACTAACAAC	900
	TTCTACCAAG	GACCCCTGGA	ACGATCCACT	GGCACTTCCA	CTAGCCTAGA	950
	GATTTCCCTC	TGGAAGACAC	TACAACTGCA	GGGCCCCCTC	TTTGGCCCTA	1000
	TCCAGCAGGA	AGTAGCTAGA	GCGGTCATCG	GCCAAATTCC	CAACAGCAGT	1050
	TGGGGTGCTC	TGTTTAGAGG	GGGGATTGAA	GAGGTGACAG	CCTGCTGGCA	1100
10	GCCTCACAGC	CCTCGTTGGY	TCTCAGTGCC	TCCTCAGCCT	TGGTGCCAC	1150
	TCTGGCCGTG	CTTGAGGAGC	CCTTCAGCCT	GCCACTGCAC	TGTGGGAGCC	1200
	TCTTTCTGGG	CTGGACAAGG	CCGGAGCCAG	CTCCCTCAGC	TTGCAGGGAG	1250
	GTATGGAGGG	AGAGATGCAG	GCGGGAACCA	GGGCTGCGCA	TGGCGCTTGC	1300
	GGGCCAGCAT	GAGTTCCAGG	TGGGCGTGGG	CTCGGCGGGC	CCCACACTCG	1350
15	GGCAGTGAGG	GGCTTAGCAC	CTGGGCCAGA	CAGATGCTGT	GCTCAACTTC	1400
	TTGCTGGGC	CTTAGCTGCC	TTCCCCGTGG	GGCAGGGCTY	CGGGAACMTG	1450
	CAGCCTGCCC	ATGCTTGAGC	CCCCCACCCC	GCCGTGGGTT	CYTGCACAGC	1500
	CCAAGCTTCC	CGGACAAGCA	CCACCCCTTA	TCCACGGTGC	CCAGTCCCAT	1550
	CAACCACCCA	AGGGTTGAGG	AGTGCGGGCA	CACAGCGCGG	GATTGGCAGG	1600
20	CAGTTCCACT	TGCGGCCCTG	GTGCGGGATC	CACTGCGTGA	AGCCAGCTGG	1650
	GCTCCTGAGT	CTGGTGGGGA	CTTGAGAAT	CTTTATGTCT	AGCTAAGGGA	1700
	TTGTAAATAC	ACCAATCAGC	AC			1722

## 25 (2) INFORMATIONS POUR LA SEQ ID NO: 59:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 495 paires de bases  
 (B) TYPE: nucléotide  
 30 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

## (ii) TYPE DE MOLECULE: ADNc

## 35 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 59:

	CTTCCCCAAC	TAATAAGGAC	CCCCCTTTCA	ACCCAAACAG	TCCAAAAGGA	CATAGACAAA	60
	GGAGTAAACA	ATGAACCAAA	GAGTGCCAAT	ATTCCCTGGT	TATGCACCCT	CCAAGCGGTG	120
	GGAGAAGAAT	TCGGCCCAGC	CAGAGTGCAT	GTACCTTTT	CTCTCTCACA	CTTGAAGCAA	180
40	ATTAAAATAG	ACNTAGGTNA	ATTNTCAGAT	AGCCCTGATG	GYTATATTGA	TGTTTACAA	240
	GGATTAGGAC	AATCCTTTGA	TCTGACATGG	AGAGATATAA	TATTACTGCT	AAATCAGACG	300
	CTAACCTCAA	ATGAGAGAAG	TGCTGCCATA	ACTGGAGCCC	GAGAGTTTGG	CAATCTCTGG	360
	TATCTCAGTC	AGGTCAATGA	TAGGATGACA	ACGGAGGAAA	GAGAACGATT	CCCCACAGGG	420
	CAGCAGGCAG	TTCCAGTGT	AGCTCCTCAT	TGGGACACAG	AATCAGAACA	TGGAGATTGG	480
45	TGCCGCAGAC	ATTTA					495

## (2) INFORMATIONS POUR LA SEQ ID NO: 60:

## 50 (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 2503 paires de bases  
 (B) TYPE: nucléotide  
 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

55

## (ii) TYPE DE MOLECULE: ADNc

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 60:

112

CCAAGAAGCC ACCAATTCCG GANCAACATTT TGGCGACCAC GAAGGGACTT TCGCATATCG 60  
 CCAAGCGGTG AGACAATAGC CGAGCGGTGA GACCTTTCCC AATCGCCAAG CAGTGAGTAC 120  
 CATCAGACCC CTTTCACTTG CTATTCTGTC CTATCTTTCT TTAGAATTCG GGGGCTAAAT 180  
 ACCGGGCATC TGTCAGCCAT TTAAAAGTGA CTAGCGGGCC GCCGGACTAA AGACACGGGT 240  
 5 GTCAAGCTTT CTGGGAAAGG GCTCTCTAAC AACCCCAAC TCTTTGGAGT TGGGACCGTT 300  
 GGTTTGCTTA GAACCAGCTT CCGCTTTTCC TGTACTTCTG GGCTGAGCCG TGGGTTGACA 360  
 GTGAAGGAAA GCCATGCATC TCCGGGGTCT CGMCAACATG TTGGTTGACC CTGCGGCCAT 420  
 GAGTGGAACT CTCAAAAGCA TGTGCCCCAA GCGACACTCG CCTATCTATC CTATCTATCC 480  
 TGACCCTTGC CCTCTGGGTC CTAATGCCTG CCAGACAAAC TTCCTCTCGC CTCTCTTCTC 540  
 10 TGAAGCTAGA ACCGCTTCTA AAAATTGCTA CTGCTCTCTT GGTGCTTTTC CTARTTTCTC 600  
 CTATAAAGAA TGAWTTCTAG TATTAAACTC CAGGACTCTG TTACCTTCTT TAGGCACCCG 660  
 GGCTCACCAA TCAGAAAGAC ACAGTTTTTG CCCAAGGCC CATCGTAGTG GGGACTACCT 720  
 GGAATTTTAT GATCCCTCCT CAGACTAACA GGCCTAACA AAGTTATTCC TGAAGCTAGG 780  
 ATATGGGGAG CCTCAGAAAT TGTATCCCTC CTATTCATAT AAGTGAGAAC AAAAGGTGTC 840  
 15 ACTCTTCCAA CCCTGAAGAT CCCCTCCCTC CCTCAGGGTA TGGCCCTCCA TTTCATTTTT 900  
 GTGGCATAAC ATCTTTATAG GATGGGGTAA AGTCCCAATA CTAACAGGAG AATGCTTAGG 960  
 ACTCTAACAG GTTTTTGAGA ATGCGTCAGT AAGGGCCACT AAATCTGATT TTTCTCAGTC 1020  
 GGTCTCTCTT GTGGTCTAGG AGGACAGGCA AGGTTGTGCA GGTTTTCGAG AATGCGTCAG 1080  
 TAAGGACCAC TAAATCCGAC CTTCTCTCGT CCTCCATGTG GTCTGGGAGG AAAACTAGTG 1140  
 20 TTTCTGCTGC TGCGTCTGTC AGCGCAACTA TTCAAGTCAG CAGGCTCCAG GGACCGTTGC 1200  
 AGGTTCTTGG GCAGGGGTTG TTTCTGCTGC TGCATTGGTG AATGCAACTA TTCTGATCAG 1260  
 CAGGGTCCCA GGACCATTCG AGGTCTTTGG GCAGGGAGAG AAACAAAACA AACCAAAACT 1320  
 GTGGGCGGTT TTGTCTTTCA TATGGGAAAC ACTCAGGCAT CAACAGGTTT ACCCTTGAAA 1380  
 TGCATCCTAA GCCATTGGGA CCAATTTGAC CCACAAACCC TGAAAAAGAG GAGGCTCATT 1440  
 25 TTTTCTGCA CTACGGCTTG GCCCAATAT TCTCTTYYTG ATGGGGAAAA ATGGCCACCT 1500  
 GAGGGAAGCA CAAATTACAA TAYTATCCTA CAGCYTGATC TTTTCTGTAA GAGGGAAGGC 1560  
 AAATGGAGTG AATACCTTAT GTCCAAGCTT TCTTTTCATT GAGGGAGAAT ACACAACTAT 1620  
 GCAAAGCTTG CAATTACAT CCCACAGGAG GACCCTTCAG CTTACCCCCA TATCCTAGCC 1680  
 TCCCTATAGC TTCCCTTCCT ATTGATGATA CTCTCTCTCT AATCTCCCTT GCCCAGAAGG 1740  
 30 AAATAAGCAA AGAAATCTCC AAAGGTCCAC AAAAACCCCC GGGCTATCGG TTATGTCCCT 1800  
 TCAAGYTGTA GGGGGAGGGG AATTTGGCCC AACCGGGGTG CATGTCCCTT CTCCCTCTCT 1860  
 GATTTAAAGC AGATCAAGGC AGACCTGGGG AAGTTTTTCAG ATGATCCTGA TAGGTACATA 1920  
 GATGTCCTAC AGGGTCTAGG GCAAACCTTT GACCTCACTT GGAGAGACGT CATGCTACTG 1980  
 TTAGATCAAA CCCTGGCCTT TAATGAAAAG AATGCGGCTT TAGCTGCAGC CTGAGAGTTT 2040  
 35 GGAGATACCT GGTATCCTAG TCAAGTAAAT GAAAGAATGA CAGCCGAAGA AAGGGACAAC 2100  
 TTCCTTACTG GTCAGCAACC CATCCCCAGT ATGGATCCCC ACTGGGACTT TGAAGGACT 2160  
 GTGAGGGACT GGAGTCGTAA ACATCTGTTG ATCTGTGTTT TGGAGGACT AAGGAGAATT 2220  
 GGGAAAAAGC CCATGAATTA TTCAATGATA TCCACCATAA CCCAGGGAAA GGAAGAAAAT 2280  
 CCTTCTGCCT TCCTCGAGCG GCTACAAGAG GCCTTAAGAA AATATACTCC CCTGTCACCC 2340  
 40 GAATCACTCG AGGGTCAATT GATTCTAAAA GATAAGTTTA TTACCCAATC AGCCACAGAT 2400  
 ATCAGGAGAA AGCTCCAAAA GCAAGCCCTG AGCCTGAACA AATCTAGAG ACATTATTAA 2460  
 ACCTGGCAAC CTGGTGTTT TATAATAGG ACCAAGAGGA ACA 2503

45 (2) INFORMATIONS POUR LA SEQ ID NO: 61:

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 1167 paires de bases  
 (B) TYPE: nucléotide  
 50 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

55 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 61:

AAGGAACTC AGAAAGCCAA TACCCATTTA GTAAGATGGA CACCAGAAGC AGAAGCAGCT 60  
 TTCCAGGCCC TAAAGAAATC CCTAACCCAA GCCCCAGTGT TAAGCTTGCC AACGGGGCAA 120  
 GACTTTTCTT TATATGTCAC AGAAAAACAG GAATAGCTCT AGGAGTCCTT ACACAGGTCC 180  
 60 AAGGGACAAG CTTGCAACCT GTGGCATACC TGAGTAAGGA AACTGATGTA NTGGCAAAGG 240  
 GTTGGCCTCA TTGTTTACAG GTAGGGCAGC AGTAGCAGTC TTAGTTTCTG AAACAGTTAA 300

113

```

AATAATACAG GGAAGAGATC TTACTGTGTG GACATCTCAT GATGTGAACG GCATACTCAC 360
TGCTAAAGAG GACTTGTGGC TGTCAGACAA CCATTTACTT AAATAGCAGG TTCTATTACT 420
TGAAGTGCCA GTGCTGCGAC TGCACATTG TGCAACTCTT AACCCAGCCA CATTCTTCC 480
AGACAAATGAA GAAAAGATAG AACATAACTG TCAACAAGTA ATTGCTCAAA CCTATGCTGC 540
5 TCGAGGGGAC CTTCTAGAGG TTCCCTTGAC TGATCCCGAC CTCAACTTGT ATACTGATGG 600
AAGTTCCTTG GCAGAAAAAG GACTTTGAAA AGCGGGGTAT GCAGTGATCA GTGATAATGG 660
AATACTTGAA AGTAATCGCC TCACTCCAGG AACTAGTGCT CACCTGGCAG AACTAATAGC 720
CCTCACTTGG GCACTAGAAT TAGGAGAAGG AAAAAGGGTA AATATATATT CAGACTCTAA 780
GTATGCTTAC CTAGTCCTCC ATGCCCATGC AGCAATATGG AGAGAGAGGG AATCCTAAC 840
10 TTCTGAGGGA ACACCTATCA ACCATCAGGG AAGCCATTAG GAGATTATTA TTGGCTGTAC 900
AGAAACCTAA AGAGGTGGCA GTCTTACACT GCCAGGGTCA TCAGGAAGAA GAGGAAAGGG 960
AAATAGAAGG CAATCGCCAA GCGGATATTG AAGCAAAAAA AGCCGCAAGG CAGGACTCTC 1020
CATTAGAAAT GCTTATAGAA GGACCCCTAG TATGGGGTAA TCCCCTCTGG GAAACCAAGC 1080
CCCAGTACTC AGCAGGAAAA ATAGAATAGG AAACCTCACA AGGACATACT TTCCTCCCCT 1140
15 CCAGATGGCT AGCCACTGAG GAAGGAA 1167

```

**(2) INFORMATIONS POUR LA SEQ ID NO: 62:**

- 20 (i) CARACTERISTIQUES DE LA SEQUENCE:
- (A) LONGUEUR: 78 paires de bases
  - (B) TYPE: nucléotide
  - (C) NOMBRE DE BRINS: simple
  - (D) CONFIGURATION: linéaire

25

- (ii) TYPE DE MOLECULE: ADNc

- (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 62:

```

30 TCCAAAGGCA CCAGGGCCCT CAGTGAGGAA CGTATCCAGC CTATACTGGC TTATCCTCAT 60
CCCAAACCC TAAAGCAA 78

```

**(2) INFORMATIONS POUR LA SEQ ID NO: 63:**

- 35 (i) CARACTERISTIQUES DE LA SEQUENCE:
- (A) LONGUEUR: 26 résidus d'acide aminé
  - (B) TYPE: acides aminés

- (ii) TYPE DE MOLECULE: peptide

40

- (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 63:

```

Ser Lys Gly Thr Arg Ala Leu Ser Glu Glu Arg Ile Gln Pro Ile Leu
1      5      10      15
45 Ala Tyr Pro His Pro Lys Thr Leu Lys Gln
      20      25

```

**(2) INFORMATIONS POUR LA SEQ ID NO: 64:**

- 50 (i) CARACTERISTIQUES DE LA SEQUENCE:
- (A) LONGUEUR: 28 paires de bases
  - (B) TYPE: nucléotide
  - (C) NOMBRE DE BRINS: simple
  - (D) CONFIGURATION: linéaire

55



114

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 64:

5 AAATGTCTGC GGCACCAATC TCCATGTT 28

(2) INFORMATIONS POUR LA SEQ ID NO: 65:

10 (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 30 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

15 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 65:

20 AAGGGGCATG GACGAGGTGG TGGCTTATTT 30

(2) INFORMATIONS POUR LA SEQ ID NO: 66:

25 (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 21 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

30 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 66:

35 GGAGAAGAGC AGCATAAGTG G 21

(2) INFORMATIONS POUR LA SEQ ID NO: 67:

40 (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 25 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

45 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 67:

50 GTGCTGATTG GTGTATTTAC AATCC 25

(2) INFORMATIONS POUR LA SEQ ID NO: 68:

115

- (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 34 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc
- (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 68:  
GACTCGCTGC AGATCGATT TTTTTTTTTT TTTT 34
- (2) INFORMATIONS POUR LA SEQ ID NO: 69:
- (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 30 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc
- (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 69:  
GCCATCAAGC CACCCAAGAA CTCTTAAGT 30
- (2) INFORMATIONS POUR LA SEQ ID NO: 70:
- (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 30 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc
- (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 70:  
CCAATAGCCA GACCATTATA TACACTAATT 30
- (2) INFORMATIONS POUR LA SEQ ID NO: 71:
- (i) CARACTERISTIQUES DE LA SEQUENCE:  
(A) LONGUEUR: 23 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire
- (ii) TYPE DE MOLECULE: ADNc

116

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 71:

GCCATAACTG CAACCCAAGA GTT

23

5

(2) INFORMATIONS POUR LA SEQ ID NO: 72:

(i) CARACTERISTIQUES DE LA SEQUENCE:

10

(A) LONGUEUR: 23 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

15

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 72:

GGACGAGGTG GTGGCTTATT TCT

23

20

(2) INFORMATIONS POUR LA SEQ ID NO: 73:

(i) CARACTERISTIQUES DE LA SEQUENCE:

25

(A) LONGUEUR: 25 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

30

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 73:

AACTTGCGTG CTAGAAGGAC TAAGG

25

35

(2) INFORMATIONS POUR LA SEQ ID NO: 74:

(i) CARACTERISTIQUES DE LA SEQUENCE:

40

(A) LONGUEUR: 24 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

45

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 74:

AACTTTTCCC TTTTCCAGAT CCTC

24

50

(2) INFORMATIONS POUR LA SEQ ID NO: 75:

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 22 paires de bases

117

- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

5 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 75:

GCATACCAGG CAAGTGGACA TT 22

10

(2) INFORMATIONS POUR LA SEQ ID NO: 76:

- 15 (i) CARACTERISTIQUES DE LA SEQUENCE:
- (A) LONGUEUR: 25 paires de bases
  - (B) TYPE: nucléotide
  - (C) NOMBRE DE BRINS: simple
  - (D) CONFIGURATION: linéaire

20 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 76:

CTGTCCGTTG GGTTCCTTA CTCCT 25

25

(2) INFORMATIONS POUR LA SEQ ID NO: 77:

- 30 (i) CARACTERISTIQUES DE LA SEQUENCE:
- (A) LONGUEUR: 24 paires de bases
  - (B) TYPE: nucléotide
  - (C) NOMBRE DE BRINS: simple
  - (D) CONFIGURATION: linéaire

35 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 77:

GAGGCTCTGG AAAAGGGAAA AGTT 24

40

(2) INFORMATIONS POUR LA SEQ ID NO: 78:

- 45 (i) CARACTERISTIQUES DE LA SEQUENCE:
- (A) LONGUEUR: 25 paires de bases
  - (B) TYPE: nucléotide
  - (C) NOMBRE DE BRINS: simple
  - (D) CONFIGURATION: linéaire

50 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: :

CTGTCCGTTG GGTTCCTTA CTCCT 25

**(2) INFORMATIONS POUR LA SEQ ID NO: 79:**

5 (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 25 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

10 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 79:

15 AGGAGTAAGG AAACCCAACG GACAG

25

**(2) INFORMATIONS POUR LA SEQ ID NO: 80:**

20 (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 25 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

25 (ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 80:

30 TGTATATAAT GGTCTGGCTA TTGGG

25

**(2) INFORMATIONS POUR LA SEQ ID NO: 81:**

35 (i) CARACTERISTIQUES DE LA SEQUENCE:  
 (A) LONGUEUR: 25 paires de bases  
 (B) TYPE: nucléotide  
 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

40 (ii) TYPE DE MOLECULE: ADNC

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 81:

45 AGGAGTAAGG AACCCCAACG GACAG

25

**(2) INFORMATION POUR LA SEQ ID NO: 82:**

50 (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 26 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

119

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 82:

5

TTCGGCAGAA ACCTGTTATG CCAAGG

26

(2) INFORMATIONS POUR LA SEQ ID NO: 83:

10

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 22 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

15

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: :

20

CTCGATTCT TGCTGGGCCT TA

22

(2) INFORMATIONS POUR LA SEQ ID NO: 84:

25

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 20 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

30

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 84:

35

GTTGATTCCC TCCTCAAGCA

20

(2) INFORMATIONS POUR LA SEQ ID NO: 85:

40

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 20 paires de bases

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

45

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 85:

50

CTCTACCAAT CAGCATGTGG

20

120

## (2) INFORMATIONS POUR LA SEQ ID NO: 86:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- 5 (A) LONGUEUR: 19 paires de bases  
 (B) TYPE: nucléotide  
 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

## (ii) TYPE DE MOLECULE: ADNc

10

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 86:

TGTTCTCTT GGTCCCTAT

19

15

## (2) INFORMATIONS POUR LA SEQ ID NO: 87:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- 20 (A) LONGUEUR: 433 résidus d'acide aminé  
 (B) TYPE: acides aminés

## (ii) TYPE DE MOLECULE: peptide

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 87:

25

25 MATATGTGIA GLSTSLSYH TLSKNFSDSL QEIMKSILT QSQDLSLAAM 50  
 TLQNRGPHL LTAEKGGCLT FLGEECCFYT NQSGIVRDAT WHLQERASDI 100  
 RQCLNSYTN LWSWATWLLP FLGPMAAILL LLTFGPCIFK LLVKFVSSRI 150  
 EAIKLQVLQ MEPQMSSTNN FYQGFLERST GTSTSLEIPL WKTLOLOGPF 200  
 30 FAPIQQEVAR AVIGQIPNSS WGVLFRRGIE EVTACWQPHS PRWXSVPQP 250  
 WCPLWPCRLS PSACHCTVGA SFWAGQGRSQ LPQLAGRYGG RDAGGNQGCA 300  
 WRLRASMSR WAWARRAPHS GSEGLSTWAR QMLCSTSSLG LSCLPRGAGL 350  
 REXAACPCLS PPPRRGFLHS PSFPDKHPL STVPSPINHP RVEECGHTAR 400  
 DWQAVFLAAL VRDPLREASW APESGGDLEN LYV 433

35

## (2) INFORMATIONS POUR LA SEQ ID NO: 88:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- 40 (A) LONGUEUR: 693 paires de bases  
 (B) TYPE: nucléotide  
 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

45

## (ii) TYPE DE MOLECULE: ADNc

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 88:

50 CTTCCCAAC TAATAAGGAC CCCCTTTCA ACCCAAACAG TCCAAAAGGA CATAGACAAA 60  
 GGAGTAAACA ATGAACCAA GAGTGCCAAT ATTCCCTGGT TATGCACCCT CCAAGCGGTG 120  
 GGAGAAGAAT TCGGCCAGC CAGAGTGCAT GTACCTTTT CTCTCTCACA CTGAGACAA 180  
 ATTAAATAG ACNTAGGTNA ATTNTCAGAT AGCCCTGATG GYTATATTGA TGTTTTACAA 240  
 GGATTAGGAC AATCCTTTGA TCTGACATGG AGAGATATAA TATTACTGCT AAATCAGACG 300  
 CTAACCTCAA ATGAGAGAAG TGCTGCCATA ACTGGAGCCC GAGAGTTTGG CAATCTCTGG 360  
 55 TATCTCAGTC AGGTCAATGA TAGGATGACA ACGGAGGAAA GAGAACGATT CCCACAGGG 420  
 CAGCAGGCAG TTCCAGTGT AGCTCCTCAT TGGGACACAG AATCAGAACA TGGAGATTGG 480

121

TGCCGCAGAC ATTTACTAAC TTGCGTGCTA GAAGGACTAA GGAAACTAG GAAGACTATG 540  
 AATTATTCAA TGATGTCCAC TATAACACAG GGGAAAGGAA GAAATCCTA CTGCCTTTCT 600  
 GGAGAGACTA AGGGAGGCAT TGAGGAAGCA TACCAGGCAA GTGGACATTG GAGGCTCTGG 660  
 AAAAGGGAAA AGTTGGGCAA ATTGAATGCC TAA 693

5

## (2) INFORMATIONS POUR LA SEQ ID NO: 89:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

10

- (A) LONGUEUR: 1577 paires de bases  
 (B) TYPE: nucléotide  
 (C) NOMBRE DE BRINS: simple  
 (D) CONFIGURATION: linéaire

15

## (ii) TYPE DE MOLECULE: ADNc

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 89:

AACTTGCGTG CTAGAAGGAC TAAGGAAAAC TAGGAAGACT ATGAATTATT CAATGATGTC 60  
 CACTATAACA CAGGGGAAAAG GAAGAAAATC CTACTGCCTT TCTGGAGAGA CTAAGGGAGG 120  
 CATTGAGGAA GCATACCAGG CAAGTGGACA TTGGAGGCTC TGGAAAAGGG AAAAGTTGGG 180  
 CAAATTGAAT GCCTAATAGG GCTTGCTTCC AGTGCAGTCT ACAAGGACGC TTTAGAAAAG 240  
 ATTGTCCAAG TAGAAATAAG CCGCCCTCG TCCATGCCCC TTATGTCAAG GGAATCACTG 300  
 GAAGGCCTAC TGCCCCAGGG GACGAAGGTC CTCTGAGTCA GAAGCCACTA ACCTGATGAT 360  
 CCAGCAGCAG GACTGAGGGT GCCCGGGGCA AGTGCCAGCC CATGCCATCA CCCTCAGAGC 420  
 CCCGGGTATG TTTGACCATT GAGAGCCAGG AAGTTAACTG TCTCCTGGAC ACTGGCGCAG 480  
 CCTTCTCAGT CTTACTTTCC TGTCCCAGAC AATTGTCTC CAGATCTGTC ACTATCCGAG 540  
 GGGTCCTAAG ACAGCCAGTC ACTACATACT TCTCTCAGCC ACTAAGTTGT GACTGGGGAA 600  
 CTTTACTCTT TTCACATGCT TTCTAATTA TGCCTGAAAG CCCCCTCCC TTGTTAGGGA 660  
 GAGACATTTT AGCAAAAGCA GGGGCCATTA TACACCTGAA CATAGGAAAA GGAATACCCA 720  
 TTTGCTGTCC CCTGCTTGAG GAAGGAATTA ATCCTGAAGT CTGGGCAATA GAAGGACAAT 780  
 ATGGACAAGC AAAGAATGCC CGTCTGTTC AAGTTAACT AAAGGATTCT GCCTCCTTTC 840  
 CCTACCAAAG GAAGTACCCT CTTAGACCCG AGGCCCTACA AGGACTCAA AGATTGTTAA 900  
 GGACCTABAA GCCCAAGGCC TAGTAAAACC ATGCAGTAGC CCCTGCAATA CTCCAATTTT 960  
 AGGAGTAAGG AAACCCAACG GACAGTGGAG GTTAGTGCAA GATCTCAGGA TTATTAATGA 1020  
 GGCTGTTTTT CCTCTATACC CAGCTGTATC TAGCCCTTAT ACTCTGCTTT CCCTAATACC 1080  
 AGAGGAAGCA GAGTAGTTTA CAGTCCTGGA CCTTAAGGAT GCCTCTTTCT GCATCCCTGT 1140  
 ACATCCTGAT TCTCAATTCT TGTGTGCTT TGAAGATCCT TTGAACCCAA TGTCTCAATT 1200  
 CACCTGGACT GTTTTACCCC AGGGGTTCG GGATAGCCCC CATCTATTTG GCCAGGCATT 1260  
 AGCCCAAGAC TTGAGCCAAT TCTCATACCT GGACATCTTG TCCTTCGGTA TGGGATGATT 1320  
 TAATTTTAGC CACCCGTTC AAGCTTGTG GCCATCAAGC CACCCAAGCG TTCTTAATT 1380  
 TCCTCACTCC GTGTGGCTAC AAGGTTTCCA AACCAAGGC TCAGCTCTGC TCACAGCAGG 1440  
 TTAATACTT AGGGTTAAAA TTATCCAAAG GCACCAGGGC CCTCTGTGAG GAATGTATCC 1500  
 AACCTGTACT GGCTTATCTT CATCCCAAAA CCCTAAAGCA ACTAAGAAGG TCCTTGGCAT 1560  
 AACAGGTTTC TGCCGAA 1577

## (2) INFORMATIONS POUR LA SEQ ID NO: 90:

50

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 182 résidus d'acide aminé  
 (B) TYPE: acides aminés

## (ii) TYPE DE MOLECULE: peptide

55

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 90:

SSSRTEGARG KCQPMPSPE PRVCLTIESQ EVNCLLDTGA AFSVLLSCPR 50  
 QLSSRSVTIR GVLRRQVTTY FSQPLSCDWG TLLFSHAFLI MPESPTPLLG 100  
 RDILAKAGAI IHLNIGKGIP ICCPLLEGI NPEVWAI EQ YGQAKNARPV 150



122

QVKLKDSASF PYQRKYPLRP EALQGLKRLR RT

182

## (2) INFORMATIONS POUR LA SEQ ID NO: 91:

5

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 36 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

10

## (ii) TYPE DE MOLECULE: ADNc

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 91:

15

AGATCTGCAG AATTCGATAT CACCCCCCCC CCCCCC

## (2) INFORMATIONS POUR LA SEQ ID NO: 92:

20

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 22 paires de bases
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION: linéaire

25

## (ii) TYPE DE MOLECULE: ADNc

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 92:

30

AGATCTGCAG AATTCGATAT CA

## REVENDICATIONS

1/ Matériel viral, à l'état isolé ou purifié, dont le génome comprend une séquence nucléotidique choisie dans le groupe incluant les séquences SEQ ID NO 46, 5 SEQ ID NO 51, SEQ ID NO 52, SEQ ID NO 53, SEQ ID NO 56, SEQ ID NO 58, SEQ ID NO 59, SEQ ID NO 60, SEQ ID NO 61, SEQ ID NO 89, leurs séquences complémentaires, et leurs séquences équivalentes, notamment les séquences nucléotidiques présentant, pour toute suite de 10 100 monomères contigus, au moins 50 % et préférentiellement au moins 70 % d'homologie avec respectivement lesdites séquences SEQ ID NO 46, SEQ ID NO 51, SEQ ID NO 52, SEQ ID NO 53, SEQ ID NO 56, SEQ ID NO 58, SEQ ID NO 59, SEQ ID NO 60, SEQ ID NO 61, 15 SEQ ID NO 89, et leurs séquences complémentaires.

2/ Matériel viral, à l'état isolé ou purifié, dont la région du génome comprenant les gènes env, pol et une partie du gène gag, à l'exclusion de la sous-région ayant une séquence identique ou équivalente à SEQ ID NO 1, 20 code pour tout polypeptide présentant, pour toute suite contigüe d'au moins 30 acides aminés, au moins 50 %, et de préférence au moins 70 % d'homologie avec une séquence peptidique codée par toute séquence nucléotidique choisie dans le groupe incluant SEQ ID NO 46, SEQ ID NO 51, 25 SEQ ID NO 52, SEQ ID NO 53, SEQ ID NO 56, SEQ ID NO 58, SEQ ID NO 59, SEQ ID NO 60, SEQ ID NO 61, SEQ ID NO 89, et leurs séquences complémentaires.

3/ Matériel viral, à l'état isolé ou purifié, dont le gène pol comprend une séquence nucléotidique 30 identique ou équivalente, partiellement ou totalement, à SEQ ID NO 57, à l'exclusion de SEQ ID NO 1.

4/ Matériel viral, à l'état isolé ou purifié, dont le gène gag comprend une séquence nucléotidique identique ou équivalente, partiellement ou totalement, à 35 SEQ ID NO 88.

5/ Matériel viral, à l'état isolé ou purifié, associé à la sclérose en plaques, dont le génome comprend une séquence nucléotidique choisie dans le groupe incluant les séquences SEQ ID NO 46, SEQ ID NO 51, SEQ ID NO 52, 5 SEQ ID NO 53, SEQ ID NO 56, SEQ ID NO 58, SEQ ID NO 59, SEQ ID NO 60, SEQ ID NO 61, SEQ ID NO 89, leurs séquences complémentaires, et leurs séquences équivalentes, notamment les séquences nucléotidiques présentant, pour toute suite de 100 monomères contigus, au moins 50 % et 10 préférentiellement au moins 70 % d'homologie avec respectivement lesdites séquences SEQ ID NO 46, SEQ ID NO 51, SEQ ID NO 52, SEQ ID NO 53, SEQ ID NO 56, SEQ ID NO 58, SEQ ID NO 59, SEQ ID NO 60, SEQ ID NO 61, SEQ ID NO 89, et leurs séquences complémentaires.

15 6/ Matériel viral, à l'état isolé ou purifié, associé à la sclérose en plaques, dont la région du génome comprenant les gènes *env*, *pol* et une partie du gène *gag*, à l'exclusion de la sous-région ayant une séquence identique ou équivalente à SEQ ID NO 1, code pour tout polypeptide 20 présentant, pour toute suite contigüe d'au moins 30 acides aminés, au moins 50 %, et de préférence au moins 70 % d'homologie avec une séquence peptidique codée par toute séquence nucléotidique choisie dans le groupe incluant SEQ ID NO 46, SEQ ID NO 51, SEQ ID NO 52, SEQ ID NO 53, 25 SEQ ID NO 56, SEQ ID NO 58, SEQ ID NO 59, SEQ ID NO 60, SEQ ID NO 61, SEQ ID NO 89, et leurs séquences complémentaires.

7/ Matériel viral, à l'état isolé ou purifié, associé à la sclérose en plaques, dont le gène *pol* 30 comprend une séquence nucléotidique identique ou équivalente, partiellement ou totalement, à SEQ ID NO 57, à l'exclusion de SEQ ID NO 1.

8/ Matériel viral, à l'état isolé ou purifié, associé à la sclérose en plaques, dont le gène *gag* 35 comprend une séquence nucléotidique identique ou équivalente, partiellement ou totalement, à SEQ ID NO 88.

9/ Matériel viral, à l'état isolé ou purifié, associé à la polyarthrite rhumatoïde, dont le génome comprend une séquence nucléotidique choisie dans le groupe incluant les séquences SEQ ID NO 51, SEQ ID NO 56, 5 SEQ ID NO 59, SEQ ID NO 60, SEQ ID NO 61, SEQ ID NO 89, leurs séquences complémentaires, et leurs séquences équivalentes, notamment les séquences nucléotidiques présentant, pour toute suite de 100 monomères contigus, au moins 50 % et préférentiellement au moins 70 % d'homologie 10 avec respectivement lesdites séquences SEQ ID NO 51, SEQ ID NO 56, SEQ ID NO 59, SEQ ID NO 60, SEQ ID NO 61, SEQ ID NO 89, et leurs séquences complémentaires.

10/ Matériel viral, à l'état isolé ou purifié, associé à la polyarthrite rhumatoïde, dont la région du 15 génome comprenant le gène *pol* et une partie du gène *gag*, à l'exclusion de la sous-région ayant une séquence identique ou équivalente à SEQ ID NO 1, code pour tout polypeptide présentant, pour toute suite contigüe d'au moins 30 acides aminés, au moins 50 %, et de préférence au moins 70 % 20 d'homologie avec une séquence peptidique codée par toute séquence nucléotidique choisie dans le groupe incluant SEQ ID NO 51, SEQ ID NO 56, SEQ ID NO 59, SEQ ID NO 60, SEQ ID NO 61, SEQ ID NO 89, et leurs séquences complémentaires.

25 11/ Matériel viral, à l'état isolé ou purifié, associé à la polyarthrite rhumatoïde, dont le gène *pol* comprend une séquence nucléotidique identique ou équivalente, partiellement ou totalement, à SEQ ID NO 57, à l'exclusion de SEQ ID NO 1.

30 12/ Matériel viral, à l'état isolé ou purifié, associé à la polyarthrite rhumatoïde, dont le gène *gag* comprend une séquence nucléotidique identique ou équivalente, partiellement ou totalement, à SEQ ID NO 88.

13/ Fragment nucléotidique comprenant une 35 séquence nucléotidique choisie dans le groupe incluant :

- (a) toutes les séquences génomiques, partielles et totales, du gène *pol* du virus MSRV-1, sauf la séquence totale du fragment nucléotidique défini par SEQ ID NO 1 ;
- 5 (b) toutes les séquences génomiques, partielles et totales, du gène *env* de MSRV-1 ;
- (c) toutes les séquences génomiques partielles du gène *gag* de MSRV-1 ;
- 10 (d) toutes les séquences génomiques chevauchant le gène *pol* et le gène *env* du virus MSRV-1, et chevauchant le gène *pol* et le gène *gag* ;
- (e) toutes les séquences, partielles et totales, d'un clone choisi dans le groupe incluant les clones FBd3 (SEQ ID NO 46), t *pol* (SEQ ID NO 51), JLBc1 (SEQ ID NO 52), JLBc2 (SEQ ID NO 53), GM3 (SEQ ID NO 56), FBd13 (SEQ ID NO 58), LB19 (SEQ ID NO 59), LTRGAG12 (SEQ ID NO 60), FP6 (SEQ ID NO 61), G+E+A (SEQ ID NO 89), à l'exclusion de toute séquence nucléotidique identique à ou comprise dans la séquence définie par SEQ ID NO 1 ;
- 20 (f) les séquences complémentaires auxdites séquences génomiques ;
- (g) les séquences équivalentes auxdites séquences (a) à (e), notamment les séquences nucléotidiques présentant, pour toute suite de 100 monomères contigus, au moins 50 % et de préférence au moins 70 % d'homologie avec lesdites séquences (a) à (e),
- 25 ledit fragment nucléotidique ne comprenant pas ou ne consistant pas en la séquence ERV-9.
- 30 14/ Fragment nucléotidique comprenant une séquence nucléotidique choisie dans le groupe incluant :
- (a) toutes les séquences génomiques, partielles et totales, du gène *pol* du virus MSRV-1, sauf la séquence totale du fragment nucléotidique défini par
- 35 SEQ ID NO 1 ;

- (b) toutes les séquences génomiques partielles du gène *gag* de MSRV-1 ;
- (c) toutes les séquences génomiques chevauchant le gène *pol* et le gène *env* du virus MSRV-1, et chevauchant le gène *pol* et le gène *gag* ;
- (d) toutes les séquences, partielles et totales, d'un clone choisi dans le groupe incluant les clones t *pol* (SEQ ID NO 51), GM3 (SEQ ID NO 56), LB19 (SEQ ID NO 59), LTRGAG12 (SEQ ID NO 60), FP6 (SEQ ID NO 61), G+E+A (SEQ ID NO 89), à l'exclusion de toute séquence nucléotidique identique à ou comprise dans la séquence définie par SEQ ID NO 1 ;
- (e) les séquences complémentaires auxdites séquences génomiques ;
- (f) les séquences équivalentes auxdites séquences (a) à (e), notamment les séquences nucléotidiques présentant, pour toute suite de 100 monomères contigus, au moins 50 % et de préférence au moins 70 % d'homologie avec lesdites séquences (a) à (e),
- ledit fragment nucléotidique ne comprenant pas ou ne consistant pas en la séquence ERV-9.

15/ Fragment selon la revendication 14, comprenant une séquence nucléotidique identique à une séquence génomique partielle ou totale du gène *pol* du virus MSRV-1, sauf à la séquence totale du fragment nucléotidique défini par SEQ ID NO 1, ou identique à toute séquence équivalente à ladite séquence génomique partielle ou totale, notamment homologue à cette dernière.

16/ Fragment nucléotidique selon la revendication 15, comprenant une séquence nucléotidique codante, partiellement ou totalement identique à une séquence nucléotidique choisie dans le groupe incluant :

- la séquence nucléotidique choisie parmi les séquences SEQ ID NO 40, SEQ ID NO 62 et SEQ ID NO 89 ;
- les séquences complémentaires à SEQ ID NO 40, SEQ ID NO 62 ou SEQ ID NO 89 ;

- les séquences équivalentes et notamment homologues à SEQ ID NO 40, SEQ ID NO 62 ou SEQ ID NO 89 ;
- les séquences codant pour tout ou partie de la séquence peptidique définie par SEQ ID NO 39, SEQ ID NO 63 ou SEQ ID NO 90 ;
- les séquences codant pour tout ou partie d'une séquence peptidique équivalente, notamment homologue à SEQ ID NO 39, SEQ ID NO 63 ou SEQ ID NO 90, susceptible d'être reconnue par des sera de patients infectés par le virus MSRV-1, ou chez lesquels le virus MSRV-1 a été réactivé.

17/ Fragment selon la revendication 13, comprenant une séquence nucléotidique identique à une séquence génomique partielle ou totale du gène env du virus MSRV-1, ou identique à toute séquence complémentaire de ladite séquence nucléotidique, ou identique à toute séquence équivalente à ladite séquence nucléotidique, notamment homologue à cette dernière.

18/ Fragment selon la revendication 13, comprenant une séquence nucléotidique identique à une séquence partielle ou totale d'un clone choisi dans le groupe incluant les clones FBd3 (SEQ ID NO 46), t pol (SEQ ID NO 51), JLBc1 (SEQ ID NO 52), JLBc2 (SEQ ID NO 53), GM3 (SEQ ID NO 56), FBd13 (SEQ ID NO 58), LB19 (SEQ ID NO 59), LTRGAG12 (SEQ ID NO 60), FP6 (SEQ ID NO 61), G+E+A (SEQ ID NO 89), à l'exclusion de toute séquence nucléotidique identique à ou comprise dans la séquence définie par SEQ ID NO 1.

19/ Fragment selon la revendication 14, comprenant une séquence nucléotidique identique à une séquence partielle ou totale d'un clone choisi dans le groupe incluant les clones t pol (SEQ ID NO 51), GM3 (SEQ ID NO 56), LB19 (SEQ ID NO 59), LTRGAG12 (SEQ ID NO 60), FP6 (SEQ ID NO 61), G+E+A (SEQ ID NO 89), à l'exclusion de toute séquence nucléotidique identique à ou comprise dans la séquence définie par SEQ ID NO 1.

20/ Sonde nucléique de détection d'un agent pathogène et/ou infectant associé à la sclérose en plaques, caractérisée en ce qu'elle est susceptible de s'hybrider spécifiquement sur tout fragment selon l'une  
5 quelconque des revendications 13 à 19, appartenant ou compris dans le génome dudit agent pathogène.

21/ Sonde nucléique de détection d'un agent pathogène et/ou infectant associé à la polyarthrite rhumatoïde, caractérisée en ce qu'elle est susceptible de  
10 s'hybrider spécifiquement sur tout fragment selon l'une quelconque des revendications 14, 15, 16, et 19, appartenant ou compris dans le génome dudit agent pathogène.

22/ Amorce pour l'amplification par polymérisation d'un ARN ou d'un ADN d'un matériel viral, caractérisée en ce qu'elle comprend une séquence nucléotidique identique ou équivalente à au moins une partie de la séquence nucléotidique d'un fragment selon l'une quelconque des revendications 13 à 19, notamment une  
20 séquence nucléotidique présentant pour toute suite de 10 monomères contigus, au moins 70 % d'homologie avec au moins ladite partie dudit fragment.

23/ Amorce selon la revendication 22, caractérisée en ce que sa séquence nucléotidique est  
25 identique à l'une quelconque des séquences choisies dans le groupe incluant SEQ ID NO 47 à SEQ ID NO 50, SEQ ID NO 55, et SEQ ID NO 64 à SEQ ID NO 86.

24/ ARN ou ADN, et notamment vecteur de répllication, comprenant un fragment génomique du matériel  
30 viral selon l'une quelconque des revendications 1 à 12, ou un fragment selon l'une quelconque des revendications 13 à 19.

25/ Peptide codé par tout cadre de lecture ouvert appartenant à un fragment nucléotidique selon l'une  
35 quelconque des revendications 13 à 19, notamment polypeptide, par exemple oligopeptide formant ou



comprenant un déterminant antigénique reconnu par des sera de patients infectés par le virus MSRV-1, et/ou chez lesquels le virus MSRV-1 a été réactivé.

26/ Polypeptide antigénique selon la  
5 revendication 25, caractérisé en ce que le cadre de lecture ouvert le codant commence, dans le sens 5'-3', au nucléotide 181, et finit au nucléotide 330 de SEQ ID NO 1.

27/ Polypeptide, notamment oligopeptide  
antigénique reconnu des sera de patients infectés par le  
10 virus MSRV-1, et/ou chez lesquels le virus MSRV-1 a été réactivé, caractérisé en ce que sa séquence peptidique est identique, partiellement ou totalement, ou équivalente à la séquence définie par SEQ ID NO 39, SEQ ID NO 63, SEQ ID NO 87.

15 28/ Oligopeptide antigénique selon la revendication 27, caractérisé en ce que sa séquence peptidique est identique à une séquence choisie dans le groupe incluant les séquences SEQ ID NO 41 à SEQ ID NO 44, SEQ ID NO 63 et SEQ ID NO 87.

20 29/ Peptide codé par tout cadre de lecture ouvert appartenant à un fragment nucléotidique selon l'une quelconque des revendications 14, 15, 16 et 19, notamment polypeptide, par exemple oligopeptide formant ou  
comprenant un déterminant antigénique reconnu par des sera  
25 de patients infectés par le virus MSRV-1, et/ou chez lesquels le virus MSRV-1 a été réactivé.

30/ Polypeptide antigénique selon la revendication 29, caractérisé en ce que le cadre de lecture ouvert le codant commence, dans le sens 5'-3', au  
30 nucléotide 181, et finit au nucléotide 330 de SEQ ID NO 1.

31/ Polypeptide, notamment oligopeptide  
antigénique reconnu des sera de patients infectés par le  
virus MSRV-1, et/ou chez lesquels le virus MSRV-1 a été réactivé, caractérisé en ce que sa séquence peptidique est  
35 identique, partiellement ou totalement, ou équivalente à la séquence définie par SEQ ID NO 39 ou SEQ ID NO 63.

32/ Oligopeptide antigénique selon la revendication 31, caractérisé en ce que sa séquence peptidique est identique à une séquence choisie dans le groupe incluant les séquences SEQ ID NO 41 à SEQ ID NO 44, et SEQ ID NO 63.

33/ Anticorps mono- ou polyclonal dirigé contre le virus MSRV-1, caractérisé en ce qu'il est obtenu par réaction immunologique d'un organisme humain ou animal, à un agent immunogène constitué par un polypeptide antigénique selon l'une quelconque des revendications 25 à 28.

34/ Anticorps mono- ou polyclonal dirigé contre le virus MSRV-1, caractérisé en ce qu'il est obtenu par réaction immunologique d'un organisme humain ou animal, à un agent immunogène constitué par un polypeptide antigénique selon l'une quelconque des revendications 29 à 32.

35/ Réactif de détection du virus MSRV-1, ou d'une exposition audit virus, caractérisé en ce qu'il comprend, à titre de substance réactive, un peptide, notamment antigénique, selon l'une quelconque des revendications 25 à 32, ou un anti-ligand, notamment anticorps dudit peptide.

36/ Composition diagnostique, prophylactique, ou thérapeutique, comprenant un peptide, notamment antigénique, selon l'une quelconque des revendications 25 à 28, ou un anti-ligand, notamment anticorps dudit peptide, selon la revendication 33.

37/ Composition diagnostique, prophylactique, ou thérapeutique, comprenant un peptide, notamment antigénique, selon l'une quelconque des revendications 29 à 32, ou un anti-ligand, notamment anticorps dudit peptide, selon la revendication 34.

38/ Composition immunothérapeutique active, notamment composition vaccinale, selon la revendication 36 ou 37.

39/ Composition diagnostique, prophylactique, ou thérapeutique, notamment pour inhiber l'expression d'au moins un agent pathogène et/ou infectant associé à la sclérose en plaques, comprenant un fragment nucléotidique selon l'une quelconque des revendications 13 à 19, ou un polynucléotide, notamment oligonucléotide, dont la séquence est partiellement identique à celle dudit fragment, sauf à celle du fragment ayant la séquence nucléotidique SEQ ID NO 1.

40/ Composition diagnostique, prophylactique, ou thérapeutique, notamment pour inhiber l'expression d'au moins un agent pathogène et/ou infectant associé à la polyarthrite rhumatoïde, comprenant un fragment nucléotidique selon l'une quelconque des revendications 14, 15, 16 et 19, ou un polynucléotide, notamment oligonucléotide, dont la séquence est partiellement identique à celle dudit fragment, sauf à celle du fragment ayant la séquence nucléotidique SEQ ID NO 1.

41/ Procédé pour détecter un agent pathologique et/ou infectant associé à la sclérose en plaques, dans un échantillon biologique, caractérisé en ce que l'on met en contact un ARN et/ou un ADN présumé appartenir ou provenant dudit agent pathologique et/ou infectant, ou leur ARN et/ou ADN complémentaire, avec une composition comprenant un fragment nucléotidique selon l'une quelconque des revendications 13 à 19, ou un polynucléotide notamment oligonucléotide dont la séquence est partiellement identique à celle dudit fragment, sauf à celle du fragment ayant la séquence nucléotidique SEQ ID NO 1.

42/ Procédé pour détecter un agent pathologique et/ou infectant associé à la polyarthrite rhumatoïde, dans un échantillon biologique, caractérisé en ce que l'on met en contact un ARN et/ou un ADN présumé appartenir ou provenant dudit agent pathologique et/ou infectant, ou leur ARN et/ou ADN complémentaire, avec une composition

comprenant un fragment nucléotidique selon l'une quelconque des revendications 14, 15, 16 et 19, ou un polynucléotide notamment oligonucléotide dont la séquence est partiellement identique à celle dudit fragment, sauf à  
5 celle du fragment ayant la séquence nucléotidique SEQ ID NO 1.

43/ Procédé pour détecter la présence ou l'exposition à un agent pathologique et/ou infectant associé à la sclérose en plaques, dans un échantillon  
10 biologique, caractérisé en ce que l'on met en contact ledit échantillon avec un peptide, notamment antigénique, selon l'une quelconque des revendications 25 à 28, ou un anti-ligand, notamment anticorps, dudit peptide, selon la revendication 33.

15 44/ Procédé pour détecter la présence ou l'exposition à un agent pathologique et/ou infectant associé à la polyarthrite rhumatoïde, dans un échantillon biologique, caractérisé en ce que l'on met en contact ledit échantillon avec un peptide, notamment antigénique,  
20 selon l'une quelconque des revendications 29 à 32, ou un anti-ligand, notamment anticorps, dudit peptide, selon la revendication 34.

45/ Dispositif de détection du virus MSRV-1, comprenant un réactif selon la revendication 35, supporté  
25 par un support solide, immunologiquement compatible avec ledit réactif, caractérisé en ce qu'il comprend un moyen de mise en contact d'un échantillon biologique, tel qu'un échantillon de sang ou de liquide céphalo-rachidien, susceptible de contenir des anticorps anti-MSRV-1, avec  
30 ledit réactif, dans des conditions permettant une éventuelle réaction immunologique, et des moyens de détection du complexe immunitaire formé avec ledit réactif.

46/ Procédé de détection d'anticorps anti-MSRV-1, dans un échantillon biologique, tel qu'un échantillon de  
35 sang ou de liquide céphalo-rachidien, caractérisé en ce qu'on met en contact ledit échantillon et un réactif selon

la revendication 35, consistant en un anticorps, dans des conditions permettant une éventuelle réaction immunologique, et on détecte ensuite la présence d'un complexe immun formé avec ledit réactif.

1/48

## FIG. 1

Consensus GTTAGGGAT ANOCTCATC TCTTGGICA GGTACTGGCC CAAGATCTAG 50  
 Consensus GGCATCTTC AGGTCCAGSN ACTCTGTTC TTGAG 85

## SEQ ID NO3 (POL MSRV-1B)

Consensus GTTAGGGAT AGOCCCCATC TATTGGGCA GGCATAGCT CAATCTGA 50  
 Consensus GGCATCTTC ATACTGGAC ATCTGTTC TTGGT 86

## SEQ ID NO4 (POL MSRV-1B)

Consensus GTTAGGGAT AGOCCCCATC TATTGGGCA GGCATAGCT CAATCTGA 50  
 Consensus GGCATCTTC ATACTGGAC ATCTGTTC TTGGT 85

## SEQ ID NO5 (POL MSRV-1B)

Consensus GTTAGGGAT AGCTCCATC TATTGGGCT GGCATTAAC CGAGACTTA 50  
 Consensus GGCATCTTC ATACTGGAC ATCTGTTC TTGG 85

## SEQ ID NO6 (POL MSRV-1B)

Consensus GTTAGGGAT AGGGGTTAR RGATANCY CATCTTTG GYARGIAT  
 Consensus RRCYRAKAY YIRGYCAVT TCTYAKRSY RGNAYCTB KYCTTYGT  
 Consensus ACATGGATGA C

## SEQ ID NO7 (POL MSRV-1B)

2/48

## FIG.2

## CONSENSUS A

SEQ ID NO 3

GTITAGGGATAGCCC TCATCTCTTGGTCA GGTACTGGCCCAAGA TCTAGGCCACTTCTC 60  
V . G . P S S L W S G T G P R S R P L L  
F R D S P H L F G Q V L A Q D L G H F S  
L G I A L I S L V R Y W P K I . A T S Q

AGGTCAGGCACTCT GTTCCTTCAG 85  
R S R H S V P S  
G P G T L F L Q  
V Q A L C S F

## CONSENSUS B

SEQ ID NO 4

GTTCAGGGATAGCCC CCATCTATTTGGCCA GGCCTAGCTCAATA CTTGAGCCAGTTCTC 60  
V Q G . P P S I W P G T S S I L E P V L  
F R D S P H L F G Q A L A Q Y L S Q F S  
S G I A P I Y L A R H . L N T . A S S H

ATACCTGGCACTCT TGTCTTCGGT 86  
I P G H S C P S  
Y L D T L V L R  
T W T L L S F G

## CONSENSUS C

SEQ ID NO 5

GTTCAGGGATAGCCC CCATCTATTTGGCCA GGCATTAGCCCAAGA CTTGAGTCAATTCTC 60  
V Q G . P P S I W P G I S P R L E S I L  
F R D S P H L F G Q A L A Q D L S Q F S  
S G I A P I Y L A R H . P K T . V N S H

ATACCTGGCACTCT TGTCTTCAG 85  
I P G H S C P S  
Y L D T L V L Q  
T W T L L S F

## CONSENSUS D

SEQ ID NO 6

GTTCAGGGATAGCTC CCATCTATTTGGCCT GGCATTAACCCGAGA CTTAAGCCAGTTCTC 60  
V Q G . L P S I W P G I N P R L K P V L  
F R D S S H L F G L A L T R D L S Q F S  
S G I A P I Y L A W H . P E T . A S S H

ATACGTGGCACTCT TGTCTTTGG 85  
I R G H S C P L  
Y V D T L V L W  
T W T L L S F

3/48

## FIG. 3

Consensus	TGGATCCAG TGYTCCACA GGGGCTGA GCTATGGG TGCAGTGGC	53
Consensus	GGATGCGGC TATAGCTCT ACGTGGATGA CCTSCGAAG CTGAG	96

SEQ ID NO 11



4/48

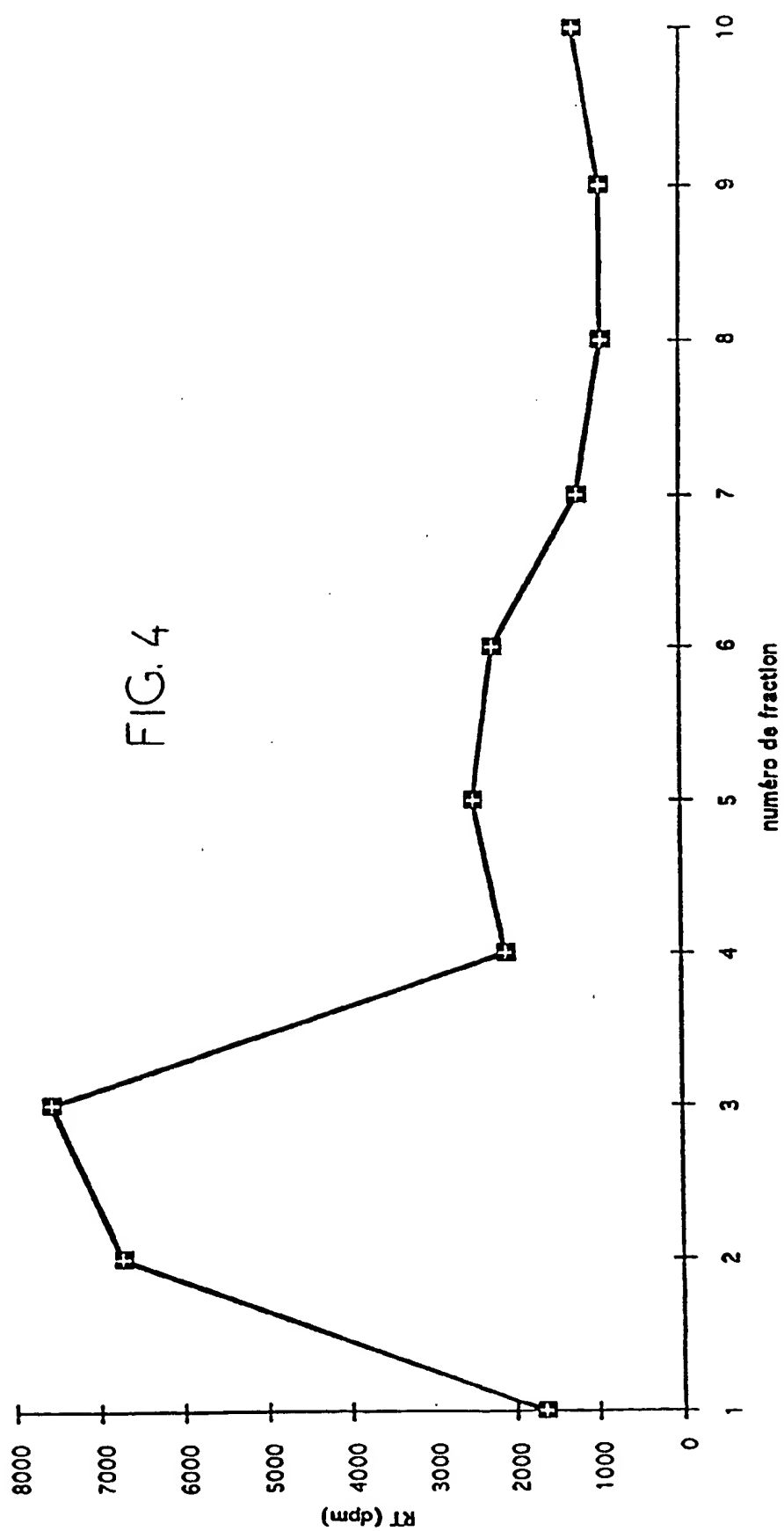
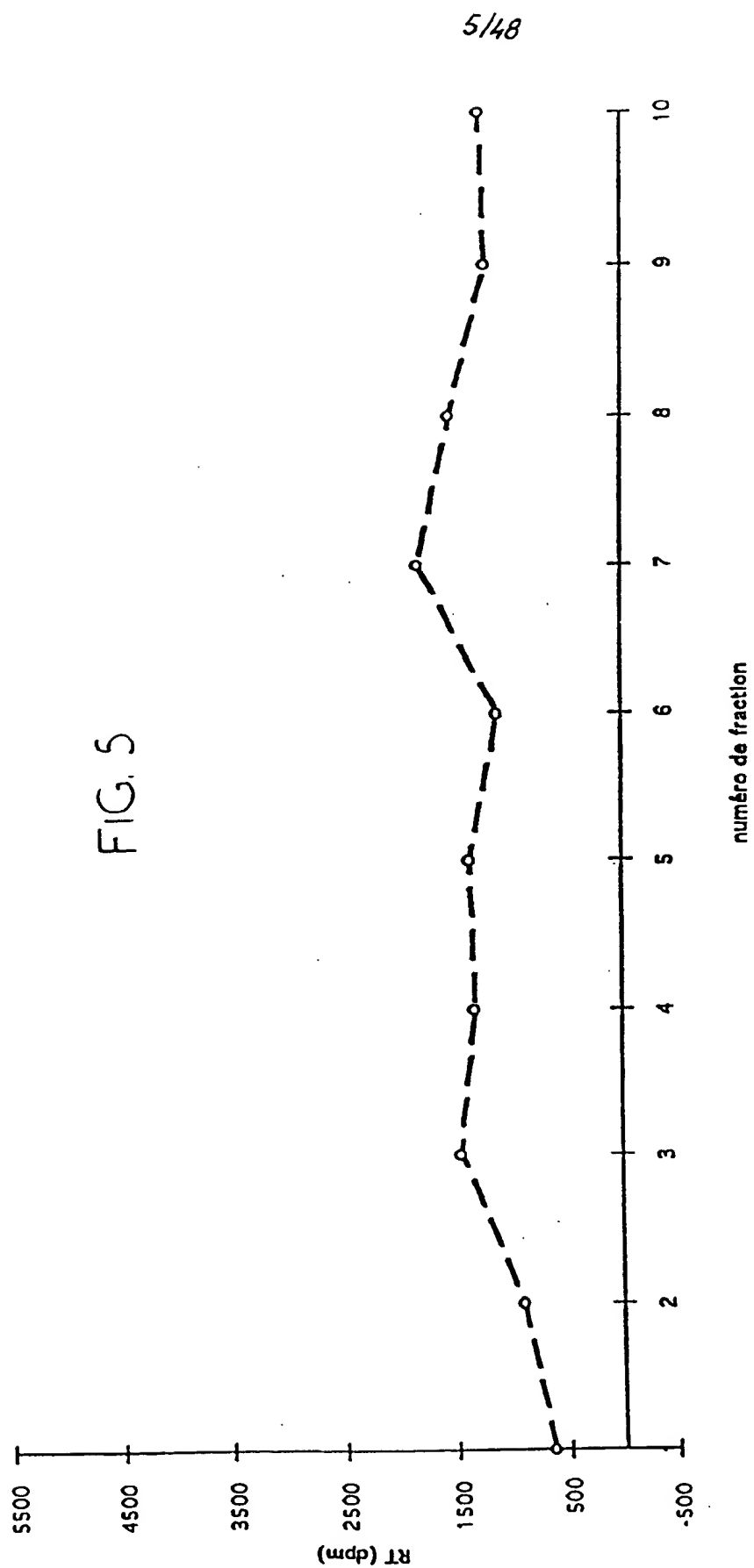


FIG. 5



6/48

## FIG. 6

CAAGCCACCC AAGAACICTT AAATTTTCCTC ACTACCTGTG GCTACAAGGT	50
TTCCAAACCA AAGGCTCAGC TCTGCTCACA GGAGATTAGA TACTTAGGGT	100
TAAAATTATC CAAAGGCACC AGGGGCTCA GTGAGGAACG TATCCAGCCT	150
ATACTGGGTT ATCCTCATCC CAAAACCTA AAGCAACTAA GAGGGTTCCT	200
TAGCATGATC AGGTTTCTGC CGAAAACAAG ATTCCAGGT ACAACCAAAA	250
TAGCCAGACC ATTATATACA CTAATTAGG AACTCAGAA AGCATTACC	300
TATTTAGTAA GATGGACACC TAAACAGAAG GCTTTCAGG CCTAAAGAA	350
GGCCTAAC CAGGCCCAG TGTTCAGCTT GGCACAGGG CAAGATTTT	400
CTTTATATGG CACAGAAAAA ACAGGAATCG CTCAGGAGT CCTTACACAG	450
GTCGAGGGA TGAGCTTGCA ACCCGTGGCA TACCTGAATA AGGAAATGA	500
TGTAGTGGCA AAGGGTGGC CTCATNGTTT ATGGGTAAATG GNGGCAGTAC	550
CAGTCINAGT ATCTGAAGCA GTTAAATTA TACAGGGAAG AGATCTTCT	600
GIGTGGACAT CTCATGATGT GAACGGCATA CTCACTGCTA AAGGAGACTT	650
GTTGTGTICA GACAACATT TACTTAANTA TCAGGCCTA TTACTTGAAG	700
AGCCAGTCT GNGACTGGC ACTTGTGCAA CTCCTAAACC C	741

SEQ ID NO9 (PSJ 17)

7/48

TCAGGGATAGCCCCCATCTATTTGGCCAGGCATTAGCCCAAGACTTGAGTC  
AATTCTCATACCTGGACACTCTTGTCTTCAGTACATGGATGATTTACTTT  
TAGTCGCCCCGTT CAGAAACCTTGTGCCATCAAGCCACCCAAGAACTCTTAA  
CTTTCCTCACTACCTGTGGCTACAAGGTTTCCAAACCAAAGGCTCGGCTCT  
GCTCACAGGAGATTAGATACTNAGGGCTAAAATTATCCAAAGGCACCAGG  
GCCCTCAGTGAGGAACGTATCCAGCCTATACTGGCTTATCCTCATCCCCAA  
ACCCTAAAGCAACTAAGAGGGTTCCTTGGCATAACAGGTTTCTGCCGAAA  
ACAGATTCCCAGGTACASCCCAATAGCCAGACCATTATATACACTAATTA  
NGGAAACTCAGAAAGCCAATACCTATTTAGTAAGATGGACACCTACAGAA  
GTGGCTTTCCAGGCCCTAAAGAAGGCCCTAACCCAAGCCCCAGTGTT CAGC  
TTGCCAACAGGGCAAGATTTTTCTTTATATGCCACAGAAAAAACAGGAAT  
AGCTCTAGGAGTCCTTACGCAGGTCTCAGGGATGAGCTTGCAACCCGTGGT  
ATACCTGAGTAAGGAAATTGATGTAGTGGCAAAGGGTT

SEQ ID NO 8 (MOO3-POO4)

FIG. 7



9/48

10 20 30 40 50 60 70  
CCC TTC GCC ACT ACA TCA ATT TTA GGA GTA AGG AA CCC AAC GGA CAG TGG AGG TTA GTG CAA GAA CTC AGG  
P F A T T S I L G V R K F N G O W R L V O E L R  
TRANSLATION OF MSRV-1 POL - (A)

80 90 100 110 120 130 140  
ATT ATC AAT GAG OCT GTT GTT OCT CTA TAC CCA OCT GTA OCT AAC OCT TAT ACA GTG CTT TCC CAA ATA OCA  
I I N E A V V P L Y P A V P N P Y T V L S O I P  
TRANSLATION OF MSRV-1 POL - (A)

150 160 170 180 190 200 210  
GAG CAA GCA CAG TGG TTT ACA GTC CTC GAC CTT AAG GAT GGC TTT TTC TCC ATC OCT GTA CGT OCT GAC TCT  
E E A E W F T V L D L K D A F F C I P V R P D S  
TRANSLATION OF MSRV-1 POL - (A)

220 230 240 250 260 270 280  
CAA TTC TTG TTT GGC TTT CAA GAT OCT TTG AAC CCA AGG TCT CAA CTC AOC TGG ACT GTT TTA OGC CAA GGC  
Q F L F A F E D P L N F T S Q L T W T V L P O G  
TRANSLATION OF MSRV-1 POL - (A)

290 300 310 320 330 340 350 360  
TTC AGG GAT AGC GGC CAT CTA TTT GGC CAG GCA TTA GGC CAA GAC TTG AGT CAA TTC TCA TAC CTC GAC ACT  
F R D S P H L F G Q A L A Q D L S Q F S Y L D T  
TRANSLATION OF MSRV-1 POL - (A)

370 380 390 400 410 420 430  
CTT GTC CTT CAG TAC ATG GAT GAT TTA CTT TTA GTC GGC OCT TCA CAA ACC TTG TCC CAT CAA GGC ACC CAA  
L V L Q Y H D D L L L V A R S E T L C H Q A T Q  
TRANSLATION OF MSRV-1 POL - (A)

440 450 460 470 480 490 500  
GAA CTC TTA ACT TTC CTC ACT ACC TGT GGC TAC AAG GTT TCC AAA CCA AAG GGT GGC CTC TCC TCA CAG CAG  
E L L T F L T T C G Y K V S K F K A R L C S O E  
TRANSLATION OF MSRV-1 POL - (A)

510 520 530 540 550 560 570  
ATT ACA TAC TTA GGC CTA AAA TTA TCC AAA GGC ACC AGG GGC CTC AGT GAG GAA GGT ATC CAG OCT ATA CTC  
I R Y X G L K L S K G T R A L S E E R I Q P I L  
TRANSLATION OF MSRV-1 POL - (A)

580 590 600 610 620 630 640  
OCT TAT GGT CAT GGC AAA ACC CTA AAG CAA CTA AGA GGC TTC CTT GGC ATA ACA GGT TTC TCC CAA AAA CAG  
A Y P H P K T L K Q L R G F L C I T G F C R K Q  
TRANSLATION OF MSRV-1 POL - (A)

650 660 670 680 690 700 710 720  
ATT GGC AGG TAC AAC CCA ATA GGC AGA CCA TTA TTT ACA CTA ATT AAG GAA ACT CAG AAA GGC AAT ACC TAT  
I P R Y X F I A R F L Y T L I X E T Q K A N T Y  
TRANSLATION OF MSRV-1 POL - (A)

730 740 750 760 770 780 790  
TTA GTA AGA TGG ACA OCT ACA GAA GTG OCT TTC CAG GGC CTA AAG AAG GGC CTA ACC CAA GGC CCA GTG TTC  
L V R W T F T E V A F Q A L K K A L T Q A P V P  
TRANSLATION OF MSRV-1 POL - (A)

800 810 820 830 840 850 860  
AOC TTG CCA ACA GGC CAA GAT TTT TCT TTA TAT GGC ACA GAA AAA ACA GCA ATA OCT CTA GCA GTC CTT AGG  
S L P T G Q D F S L Y A T E K T G I A L G V L T  
TRANSLATION OF MSRV-1 POL - (A)

870 880 890 900 910 920 930  
CAG GTC TCA GGC ATG AOC TTG CAA GGC GTG GAA TAC CTG AGT AAG GAA ATT GAT GTA CTG CCA AAG GGT TGG  
Q V S G H S L Q P V V Y L S K E I D V V A K G W  
TRANSLATION OF MSRV-1 POL - (A)

940 950 960 970 980 990 1000  
OCT CAT AGT TTA TGG GTA ATG CAG GCA GTA GCA GTC TTA GTA TCT GAA GCA GTT AAA ATA ATA CAG GGA AGA  
P H X L W V H X A V A V X V S E A V K I I O G R  
TRANSLATION OF MSRV-1 POL - (A)

1010 1020 1030 1040 1050 1060 1070 1080  
GAT CTT NCT GTG TGG ACA TCT CAT GAT GTG AAC GGC ATA CTC ACT OCT AAA GCA GAC TTG TGG TTG TCA GAC  
D L X V W T S H D V N G I L T A K G D L W L S D  
TRANSLATION OF MSRV-1 POL - (A)

1090 1100 1110 1120 1130 1140 1150  
AAC CAT TTA CTT AAT TAT CAG OCT CTA TTA CTT GAA CAG CCA GTG CTG NGA CTG GGC ACT TGT CCA ACT CTT  
N H L L X Y Q A L L L E E P V L X L R T C A T L  
TRANSLATION OF MSRV-1 POL - (A)

AAA CCC  
K P  
TRANSLATION OF MSRV-1 POL - (A)

FIG. 9

SEQ ID NO 1 (MSRV-1 pol.)

10/48

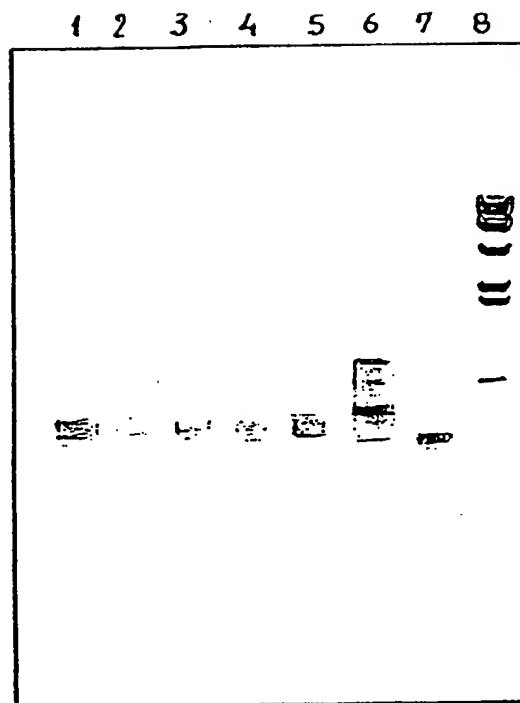
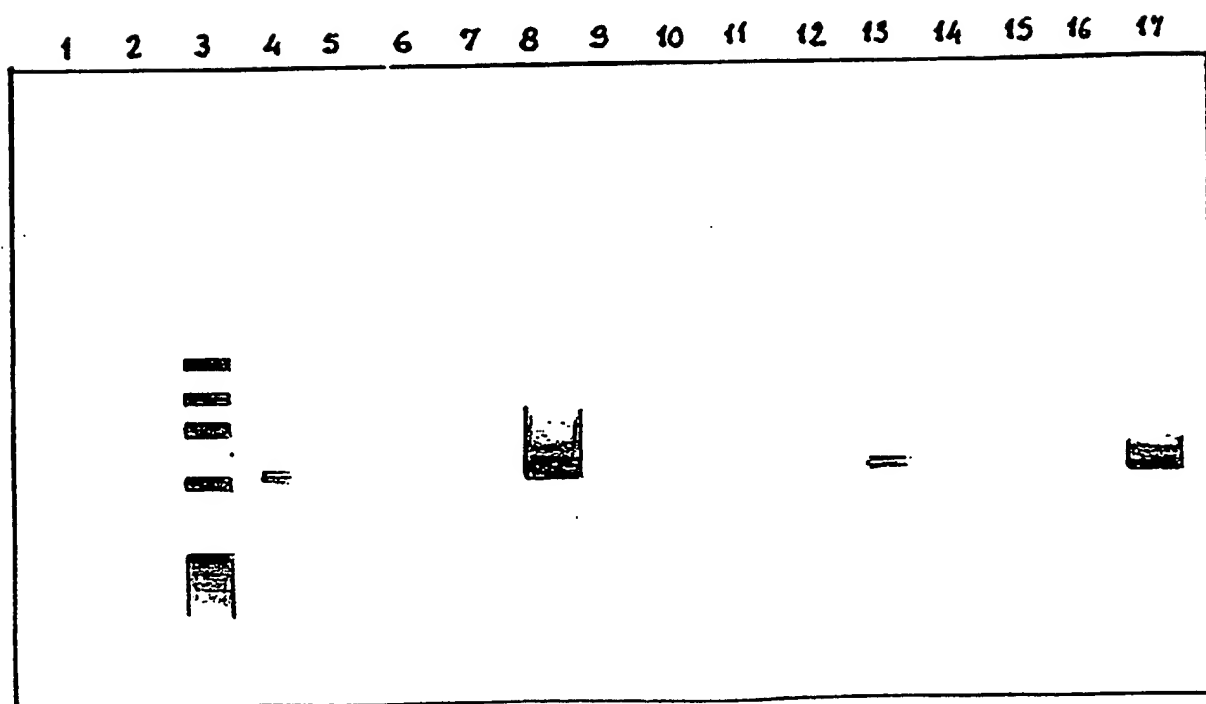


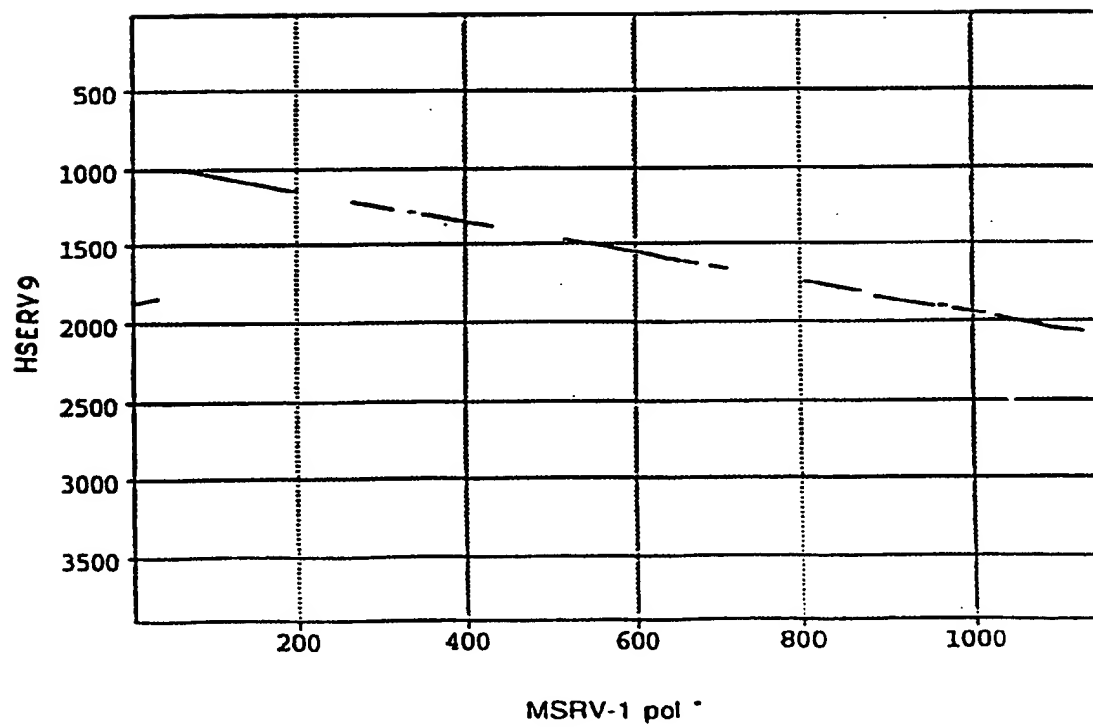
FIG. 10

FIG. 11



11/48

FIG. 12





12/48

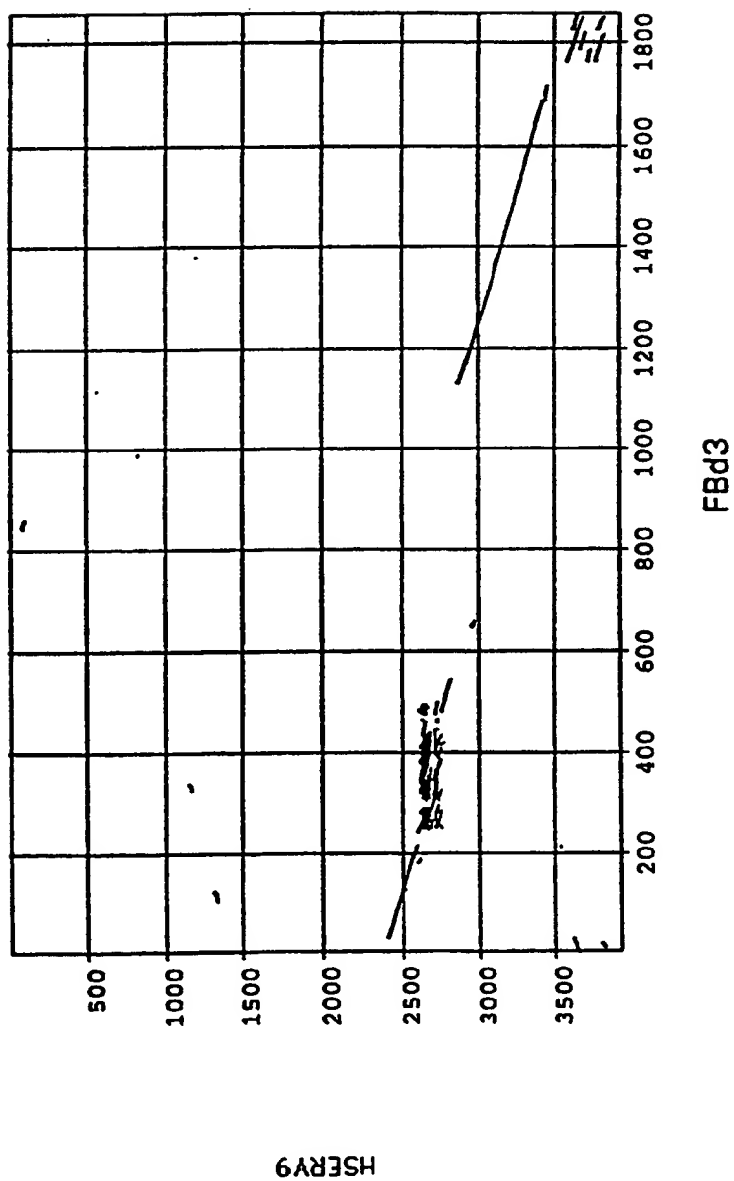
FIG. 13

SEQ ID NO 46 (FBd3)

GTGCTGATTGGTGTATTTACAATCCTTTATCTAATCCGAAATGCCCATGTTG  
CAATATGGAAAGAAAGGGAGTTCCTAACCTCTGGGGGAACCCCATTTAAA  
TACCACAAGTAAATCATGGAGTTATTGCACACAGTGCAAAAACCTCAAGGA  
GGTGGAAGTCTTACACTGCCAAAGCCATCAGAAAAGGGAAGAGGGGAGAA  
GAGCAGCATAAGTGGCTACAGAGGCAAGGAAAGACTAGCAGAAAGGAAA  
GAGAGAAAGAGACAGAAAGTCAGAGAGAGAGAGAGGAAGAGACAGAGCA  
CAAAGAGGGGAGTCAGAGAGAGAGAGAGACAGAGAGTCAGAGAGAAGGAA  
AGAGAGAGAGGAAGAGACAAAGAATGAATCAAACAGAGAGACAGAAAGT  
CAGAGAGAGAGAGAGAGAGGAAGAGACAGAGAAAAAGAGGGAGTCAGAA  
AAAGAGAGACCAAAGAAGAAGTCCAAAGAGAAAGAAAGAGAGATGGAAG  
TAGTAAAGGAAAAACAGTGTACCCTATTCTTTAAAAGCCGGGGTAAATTT  
AAAACCTATAATTGATAACTGAAGGTCTTCTCTGTAACCCTGTAACACTCC  
AATACCACCTTGTTGTCAAGTGTAACAAGGGCGTAGCCCAAAGCACTG  
AGGCCACTAACAACCCATAGCCTTCCTATCAAAATTCCTTAACCCAGCAGG  
TTTCCTAACAGGGGATCTAAATCTTAATTAATTACCATACAATGGTCCAAC  
CAGACTTAGGAGGAATTCCCTTCAGGACGGGAAGATAGATGCTTCCTCCCA  
GGCGATTAAGGGAGAAAGACACAATGGGTATTCAGTAAGTGCCAAGGGGA  
ACACTTGTTAGAAGCAAAGTTAGGAAAATTGCCAAATAATTGGTTTGCTCAA  
GAGTTGTTTGCACTCAGCCAAACCTTGAAGTACTTGCAGAATCAGAAAGGA  
GCCATCTATACCAATTCTAAGTTAATATGGACTGAAGGAGGTTTTATTAAAT  
ACCAAAGAGAAATTAATAATCCCAAACCTTATAAGGTTTTCAACCAAAGTAA  
AGTTTGCTAAAAGTTAACAGCGTAACATGTATTATCCTACTACCACACACT  
CTCAAAGGATTTCTCAGACAGTTTGCAAGAAATAATGATATCTATCCTTAC  
TCTACAATCCCAAATAGACTCTTTGGCAGCAGTGACTCTCCAAAACCGTCA  
AGGCCTAGACCTCCTCACTGCTGAGAAAGGAGGACTCTGCACCTTCTTAAG  
GGAAGAGTGTGTCTTTACACTAACCAGTCAGGGATAGTATGAGATGCTGC  
CCGGCATTTACAGAAAAAGGCTTCTGAAATCAGACAACGCCTTTCAAATTC  
CTATACCAACCTCTGGAGTTGGGCAACATGGTTTTCTTCCCTTTCTATGTCCC  
ATGGCTGCCATCTTGCTATTACTCGCCTTTGGGCCCTGTATTTTTAACCTCC  
TTGTCAAATTTGTTTCTTCTAGGATCGAGGCCATCAAGCTACAGATGGTCTT  
ACAAATGGAACCCCAAATGAGCTCAACTATCAACTTCTACTGAGGACCCCT  
AGACCAACCCCTGGCCCTTTCACTGGCCTAAAGAGTTCCCGTCTGGAGGA  
CACTACCACTGCAGGGGCCCATCTTTGCCCTATCCAGAAGGAAGTAGCTA  
GAGCAGTCATTGCCCAATTCCTAAGAGCAGCTGGGGTGTCCCGTTTAGAGT  
GGGGATTGAGAGGTGAAGCCAGCTGGACTTCTGGGTCTGGGTGGGGACTTG  
GAGAACTTTTGTGTCTAGCTAAAGGATTGTAATGCAACAATCAGTGCTCT  
GTGTCTAGCTAAAGGATTGTAATACACCAATCAGCAC

13/48

FIG. 14



14/48

## FIG. 15

SEQ ID NO 51 ( t pol)

GGCTGCTAAAGGAGACTTGTGGTTGTCAGACAATCGCCTACTTAGGTACCA  
GGCCTTATTACTTGAGGGACTGGTGCTTCAGATGCGCACTTGTGCAGCTCT  
TAACCCAAACTTATGCTGCCCAGAAGGATCTTTTAGAGGTCCCCTTAGCCA  
ACCCTGACCTCAACCTATATATACTGATGGAAGTTCGTTTGTAGAAAAG  
GGATTACAAAGGGNAGGATATNCCATAGGTTAGTGATAAAGCAGTACTTG  
AAAGTAAGCCTCTTCCCCCAGGGACCAGCGCCCCCGTTAGCAGAACTAGT  
GGCACTGACCCCGAGCCTTAGAACTTGGAAGGGAGGAGGATAAATGTGT  
ATACAGATAGCAAGTATGCTTATCTAATCCGAAATGCCCATGTTG

15/48

SEQ ID NO 52 (JLBc1)

TCAGGGATAGCCCCCATCTATTTGGTCAGGCACTGGCCCAAGATCTAGGGA  
CATGCCACTTTTAAGAGCCATTTCTCAAGTCCAGGTACTCTGGTCCITCGGT  
ATGTGGATGATTTACTTTTGGCTACCAAGTTCAGTAGCCTCATGCCAGCAGG  
CTACTCTAGATCTCTTGAACCTTTCTAGCTAATCAAGGGTACAAGGCATCTA  
GGTTGAAGGCCAGCTTTGCCTACAGCAGGTCAAATATCTAGGCCTAATCT  
TAGCCAGAGGGACCAGGGCACTCAGCAAGGAACAAATACAGCCTATACTG  
GCTTATCCTCACCTAAGACATTAAACAGTTGCGGGGGTTCCCTTGAATC  
ACTGGCTTTTTGGTGACTATGGATTCCCAGATACAGCAAGATTGGCAGGCC  
CCTCTATACTGTAATCAAGGAGACTCAGGAGGGCAAGTACTCATCTAGTAG  
AATGGGAAGTACAGGACAGAAACAGCCTTCAAAACCTTAAAGCAGGCCCTA  
GTACAATCTCCAGCTTTAAGCCTTCCCACAGGACAAAACCTTCTCTTTATAC  
ATCACAGAGAGGGCAGAGATAGCTCTTGGTGTCTTATTTCAGACTCATGGG  
ACTACCCCAACACAGTGGCACACCTAAGTAAGGAAATTGATGTAGTAGC  
AAAAGGCTGGCCTCACTGTTTATGGGTAGCTGTGGTGGTGGCTGTCTTAGT  
GTCAGAAGCTATCAAAATAATACAAGGAAAGGATCTCACTGTCTGGACTA  
CTCATGTGTAATGGCATACTAGGTGCCAAAAGAAGTTTATGGGTATCAGA  
CAACCACCTGCTTAGATACCAGGGACTACTCCTGGAGGATTGGGCTTCAAG  
TGCGTTTTTTGTGGCCTCAACCCTGCCACTTTTCTCCAGAGGATGGAGAG  
CCGCTTGAGCATGCTTGCCAACAGGTTGTAGGCCAGAATTATCCACCCGA  
GATGATCTCTTAGAGTACCCTTAGCTAATCCTGACCTTAACCTATATACCA  
ATGGAAGTTCATTTGTGGAAAACGGGATATGAAGGGCAGGTTATGTCATAG  
TTAGTGATGTAATCATACTTGCAAGTAAGCCTCTTACCCCAAGGGGCCAGCA  
CTCAGTTAGCAGAACTAGTCACACTTACCTTAACCTTAGAACTGGGAAAGG  
GAAAAAGAATAAATATGTATACAGATAGTAAGTATGCTTATCTAATCCTAC  
ATGCCCATGCTGCAATATGGAAGGAAAGGGAGTTCCTAACCCTGGGGGA  
ACCCCATTAATAACCAAGGYAAATCATGGAGTTATTGCACGCAGTGC  
AAAACTCAAGGAGGTGGCAGTCTTACACTGCCGAAGCYATCAAAAAGGG  
GAAGGAGAGGGGAGAACAGCAGCATAAGTGGTTGGCAGAGGCAGTGAAA  
GACCAGCAGAGAGAAGGAGAGAGACAACGTCAACGACAGAAGGAAAGAA  
GAGGAGGAGACAGAGAGGAAGAGACAGAGAGACAGTTAGTCCAAGAGAG  
AGACAGAGAGAGGAAGAGACAGACAGAAAGTCCAAGAGAGAAGGAAAGA  
GAGGAAGAGACCAAGGAGTCCNAGAGAGAGAAAGAGATAGAAGTAGTAA  
AGAAAAAACATTGTACCTATTTCCTTTAAAAGCCGGGGTATATTTAAAACC  
TATAATTGATAATTGAGTTCTTGCACCCTCCTCCAGGGGATYGCTGGGAGG  
AAACCCTCAACCGATATGTGAAAATTGTGGGTCGTCCCTATGTCTCAATTA  
CCAGCCAATACCCCTTGTTTTTAGTGTGAACGAGGGTGTAGAGCGCAGAC  
AGGGAGACCTCTGACAATCCATACCCTTCCTATCCAAAATCCTTAACCCAG  
CAGGTTTTCTAAAAGGGGATCTAAATCTTAATTAATTACCATACAAAGGTC  
AAACCAGATCTAGGAGGAACTTCCTTCAGGACAGGATGATAGATGGTTCCT  
CCCAGGCGATTAAAGAAAATAAAAAGACACATGGGCAGCCAGTAAGTGAT  
AAGGGAACACTAGTAGAAGCAGTTAGGAGAAGTTGCCTAATAATTGGTCT  
ACTCCAAATGTGTGAGTTGTTTCGCACTCAGCCCAAATCTTAAAGTACTTAC  
AGAATTAGGGAGGAGCCATTTACACCAATTCTAAGTTAATATGGACTGGAT  
GAGGTTTTATTAATAGCGAAGGAGAATTAATCCTAAACTNACAAGGTTTT  
CAACTAAAGTAAATTTTACTAAAAGCTAACAGTGTAACATGCATTATCCTA  
CTACAACACACTCTCANAGGATTCCTCAGACAGTTTACAAGAAATAACAA  
AATCTATCTGGTAAGGATAGTAATACTACAATCCCAAATACATTCTTTGGCAG  
CAGTGACTCTC

FIG. 16

16/48

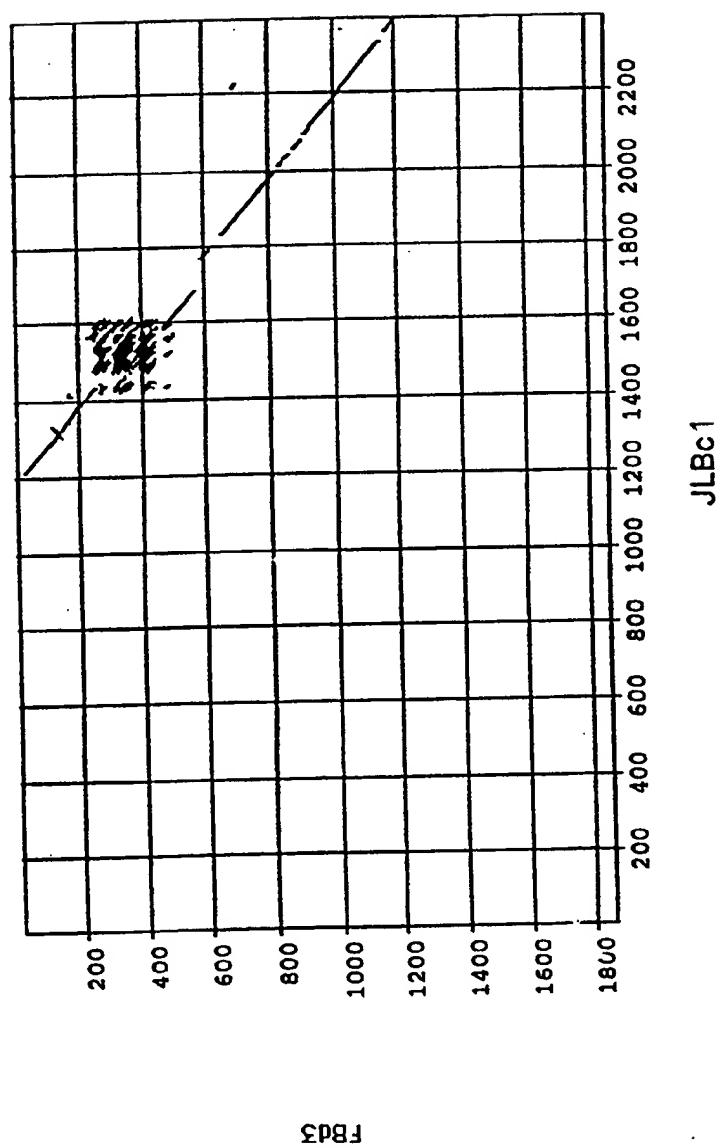
SEQ ID NO 53 (JLBc2)

TCAGGGATAGCCCCC.ATCTATTTGATCAGGCACTAGCCCAAGATCTAGGCC  
ACTTCTGAAGTCCAGGCATTCTAGTCCTTCAGTATGTGGATGATTTACTTTT  
GGCTACCAAGTTTGGGAAGCCTCATGCCAGCAGGCTACTTGAGATCTCTTGAA  
CTTTCTAGCTAATCAAGGGTGTATGGCATCTAAATTGAAAGTCCAGCTCTG  
CCTACAACAAGTCAAATATCTAGGCCTAATCTTAGATAGAAGAACCAGGG  
CCCTCAGCAAGGAATGAATAAAGCCTATGCTGGCTTATCGGCACCCTAAGA  
CATTAAAAACAATTGTGGGGGTTCCTTGGGAATCACTGGCTTTTGCCGACTAT  
GGATCCCTGGATAGAGTGAGATAGCCAGGCCCCCTCTATTACTCTTATCAA  
GGAGACCCAGAGGGCAAATACTTATCTAGTATTATGGGNACCAGAGGCAG  
AAAAAGCCTTCCAAACCTTAAAGGAGACCCTAGTACAAGCTCCAGCTTTAA  
GCCTTCCCACAGGACAAANCTTCTCTTTATATGTCACAGAGAGAGCAGGAA  
TAGCTCCTGGAGTCCTTACTCAGACTTTTGGACGACCCACGGCCAGTGGC  
RTACCTAAGTAAGGAAAATTGATGTAGTAGCAAAAGGCTGGCCTCACTGTTT  
ATGGGTAGTTGCGGCTGTGGCAGTCTTACTGTCAAAGGCTATCAAAATAAT  
ACAAGGAAAGGATTTCACTATCTGGACTACTCATGAGGAAAATGGCATATT  
AGGTGCCAAAGGAAGTTTGGCTATCAGACAACCACCTGCTCAGATTCCA  
GGCACTACTGATTGAGAGACCAGTGCTTTAAATATGTATGTGTGTGTGG  
CCCTCAACCCTGCCACTGTTCTCCAGAAAGATGGAGAACCAATGAAGCATT  
ACTGTCAACAAATTAGAGTCCAGAGTTATGCTGCCTGAGAGGATCTCTTAG  
AAGTCCCTTAGCTAATCCTGACCTTAACCTATATGCTGATGGAAGTTCAC  
TTGTGGAGAATGGGATACGAAAAGCACATTATGCCATAGTTAGTGAGGTA  
ACAGTACTTGAAAGTAAGCCTATTCCCCCATGGACCAGAGCCAGTTAGCA  
GAACTAGTGGCACTTACCCAAGCCTTAGAACTAGGAAAGGGAAAAATAAT  
AAATGTGTATACAGATAGCAAGTATGCTTATCTAATCCTACATGCCCATGC  
TGCAGTATGGAAGAAAGGGAGTTCTTAACCTCTGGGGGAACCCCCATTA  
AATACCACAAGGCAAATCATGGAGTTATTGCATGTAGTGCAAAACCTCAA  
GTAGGTGGCAGTTTTACACTGCCTGAAGCTATGGGGAAGGAGAGAGGAGA  
ACAGCAGCATAAGTGGCTAGCAGAGGCAGCGAAAGACTAGCAGAGAGGA  
GAGGTAGGGGAAAGACAGAAAGTCAAAGAAAAGAAGTCAAAGACAGACA  
GAGAAAGAGACAGAGGGAGCCAGAGAGAAAGAAAAGAGAGAACGAAAGA  
GACAGAATGTCAAAGAACAGAAGAGAGAGGCGAGCGCCAGAAGAGTTAAG  
AAAGTGAGAAAGAGAGATGGAAATAGTAAAGAAAAAACAGTGTACCCTAT  
TCCTTTAAAAGCCAGGGTAAATTTAAAACGTATAATTTTATAATTGGAAGG  
TCTTCTCCATAACCCTATAACATTAATAACACCTTGTTGTGTCAGTGTA AAC  
AAGAGCATAGCCCAAAGCACTGAGGCCACTGACAACCCATAGCCTTCCT  
ATCAAAAATCCTTAACCTCTGCAGGTTTCCTAACAGGGGATCTAAATCTCAA  
CTAATCACCATAACAATGGTCCGACCAGACCTAGGAGCGACTCCCCTCAGG  
ACAGAAGGATGGATGGTTCCCTCCAGGCCATTAAGGGAAAGAGACACAAT  
GGGTATTCAGTAAGTGATAAGGGAACCTCTGTAGAAGCAGTTAGGAAGATT  
GCCTAATATTTGGTCTGCTCAAATGTGCCAGCTGTTTGCACTCAGCTAAAC  
CTTAAATTACTTACAGAATTAGGAAGGAGCCATCTATACCAATTCTGAGTT  
AATATGAGCTGAACAAGTTCTTATTAATAGCAAAGAATCATTGAAATCTCA  
AACTTGCAAAGTTTTCAACAAAAGTAAAGTTTGCTGAAAGTTAGCAGTGTA  
ACATGTATTATCCTAACTTCTAATCTGTGGAAATCAGACCCTATCAGTGC  
CCCTCAAAGCTGAAGTCCATCAGCATATGGCCATACAATAATACCCCTAT  
TTATAGGGTTAGGAATGGCCACTGCTACAGGAATGGGAGTAACAGGTTTAT  
CTACTTCATTATCCTATTACCACACACTCTTAAAGGATTTCTCAGACAGTTT  
ACAAGAAATAACAAAATCTATCCTTACTCTINTARTCCCAAATAGRTTCTTT  
GGCAGCAGTGACTCTC

FIG. 17

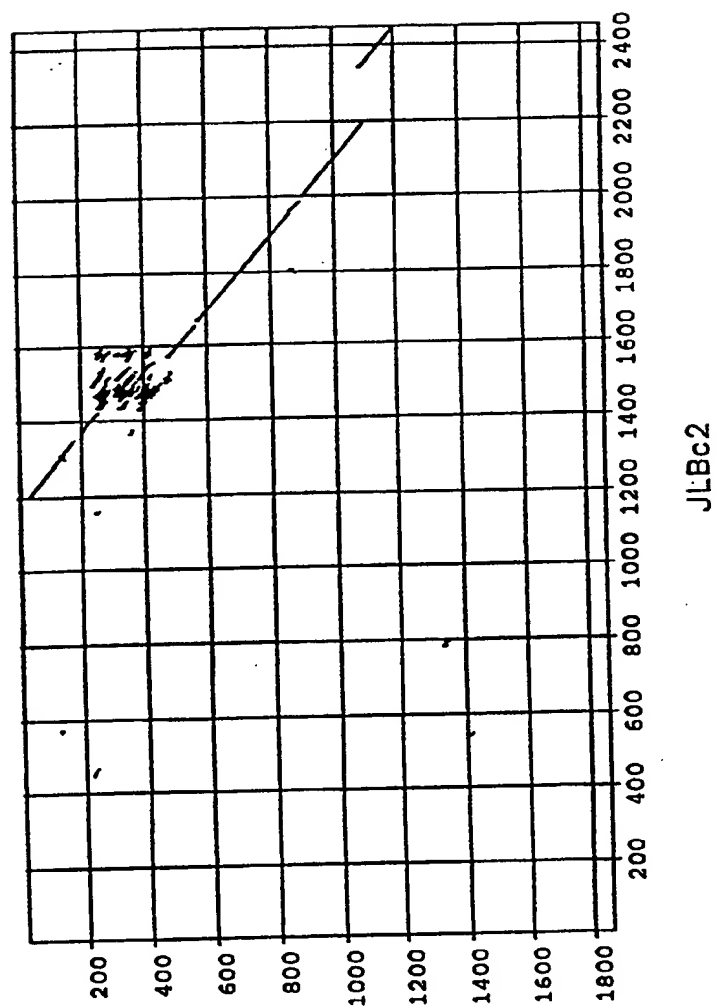
17/48

FIG. 18



18/48

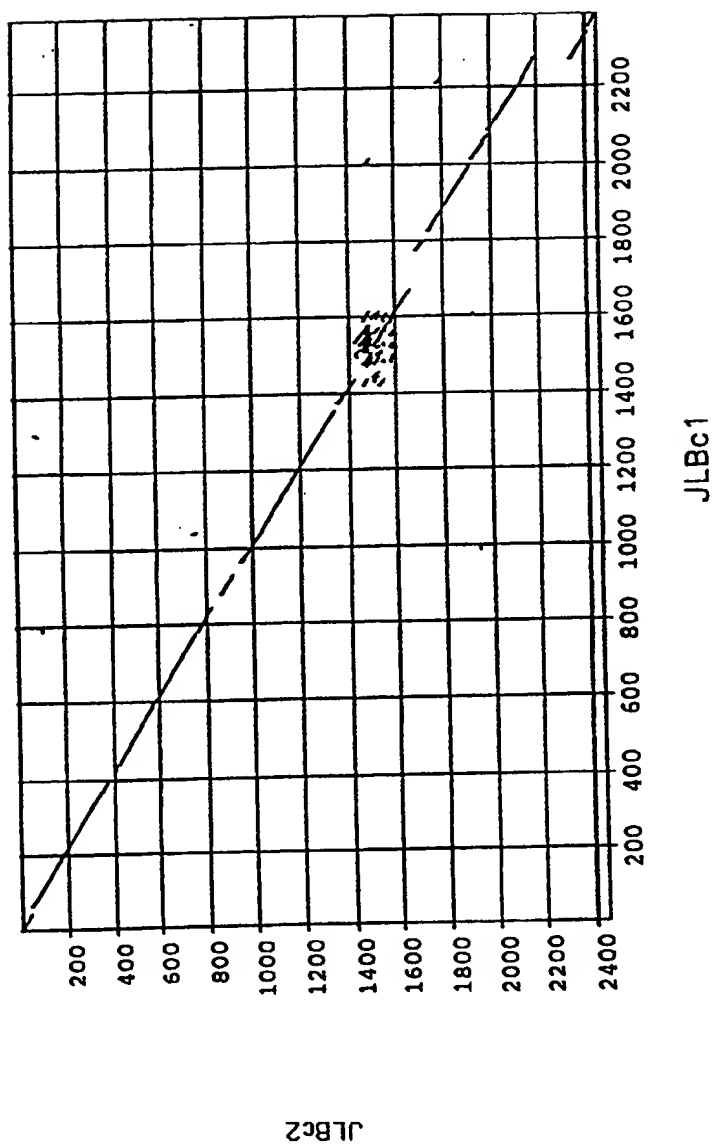
FIG. 19



FBd3

19/48

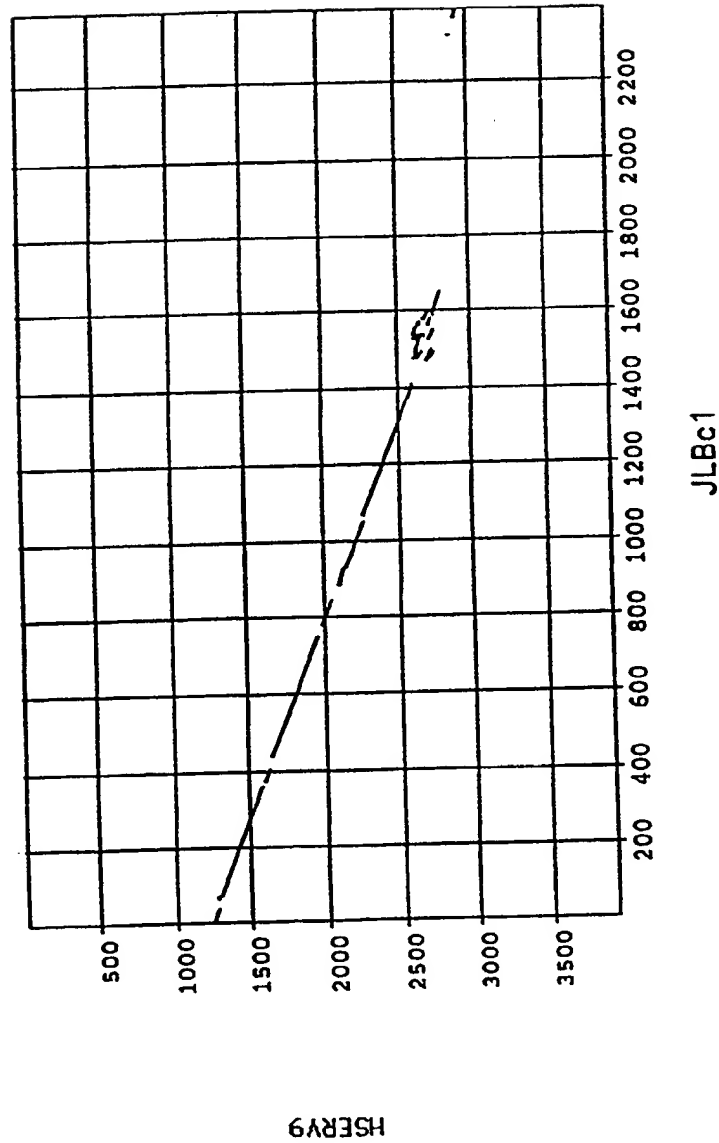
FIG. 20





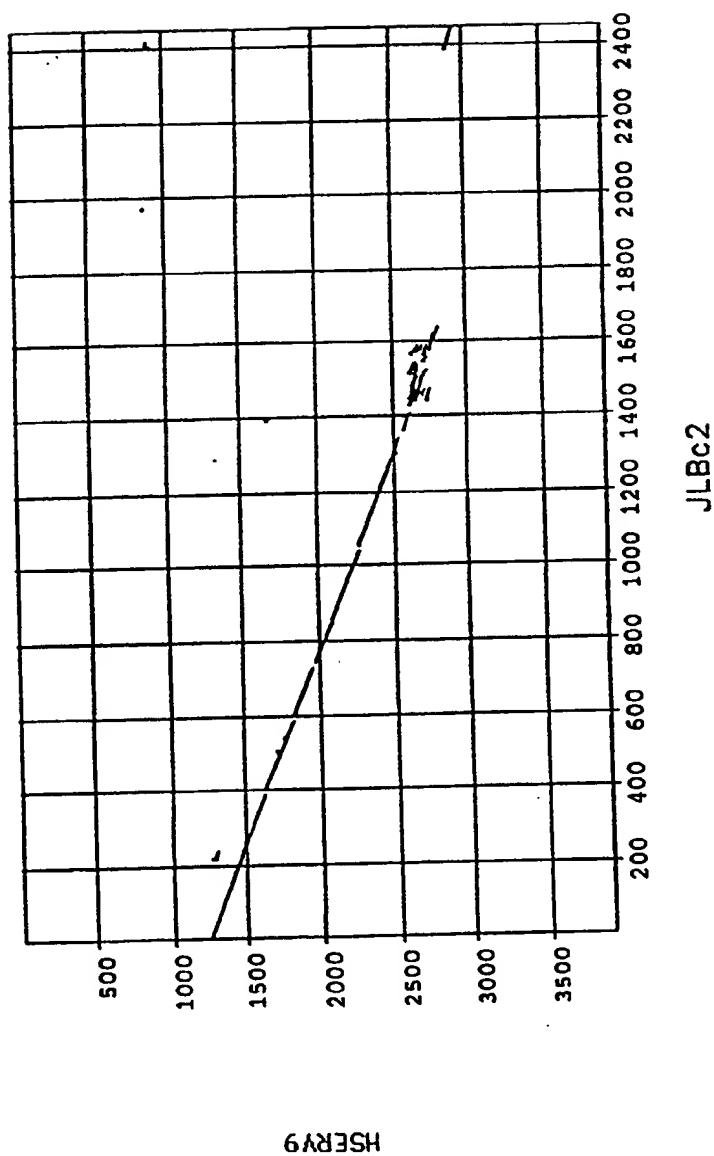
20/48

FIG. 21



21/48

FIG. 22



22/48

```
1  TTCCTGAGTT  CTTGCACTAA  CCTCAAATGA  GAGAAGTGCC  GCCATAACTG  CAAACCAAGA
61  GTTTGGCGAT  CCCTGGTATC  TCAGTCAGGT  CAATGACAGG  ATGACAACAG  AGGAAAGATA
121  ATGATTCCCC  ACAGGCCAGC  AGGCAGTTCC  CAGTGTAGAC  CCTCATTAGG  ACACAGAATC
181  AGAACATGGA  GATTGGTGCC  GCAGACATTT  GCTAACTTGC  GTGCTAGAAG  GACTAAGGAA
241  AACTAGGAAG  ATATGAATTA  TTCAATGATG  TCCACTATAA  CACAGGGGAA  AGGAAGAAAA
301  TCCTACTGCC  TTTCTGGAGA  GACTAAGGGA  GGCATTGAGG  AAGCATACCA  GGCAAGTGGA
361  CATTGGAGGC  TCTGGAAAAG  GGAAAAGTTG  GGAAAAGTAT  ATGTC TAATA  GGGCTTGCTT
421  CCAGTGTGGT  CTACAAGGAC  ACTTTAAAAA  AGATTGTCCA  ATAGAAATAA  GCCACCACCT
481  CGTCCATGCC  CCTTATGTCA  AGGGAATCAC  TGGAAGGCCC  ACTGCCCCAG  GGGATGAAGG
541  TCCTCTGAGT  CAGAAGCCAC  TAACCAGATG  ATCCAGCAGC  AGGACTGAGG  GTGCCCCGGG
601  CAAGCGCCAG  CCCATGCCAT  CACCCTCACA  GAGCCCCAGG  TATGCTTGAC  CATTGAGGGT
661  CAGAAGGGTA  CTGTCTCCTG  GACACTGGCG  GGCCTTCTCA  GTCTTACTTT  CCTGTCCTGG
721  ACAACTGTCC  TCCAGATCTG  TCACTGTCCG  AGGGGTCTTA  GGACAGCCAG  TCACTAGATA
781  CTTCTCCCAG  CCACTAAGTT  GTGACTGGGG  AACTTTACTC  TTCCACATGC  TTTTCTAATT
841  ATGCTGAAA  GCCCCACTCT  CTTGTTAGGG  GAGAGACATT  CTAGCAAAAG  CAGGGGCCAT
901  TATACATGTG  AATATAGGAG  AAGGAACAAC  TGTTTGTTGT  CCCCTGCTTG  AGGAAGGAAT
961  TAATCCTGAA  GTCCGGGCAA  CAGAAGGACA  ATATGGACAA  GCAAAGAATG  CCCGTCCTGT
1021  TCAAGTTAAA  CTAAAGGATT  CCACCTCCTT  TCCCTACCAA  AGGCAGTACC  CCCTCAGACC
1081  CGAGACCCAA  CAAGAACTCC  AAAAGATTGT  AAAGGACCTA  AAAGCCCAAG  GCCTAGTAAA
1141  ACCAAGCAAT  AGCCCTTGCA  AGACTCCAAT  TTAGGAGTA  AGGAAACCCA  ACGGAC
```

SEQ ID NO 56 (GM3)

FIG. 23

23/48

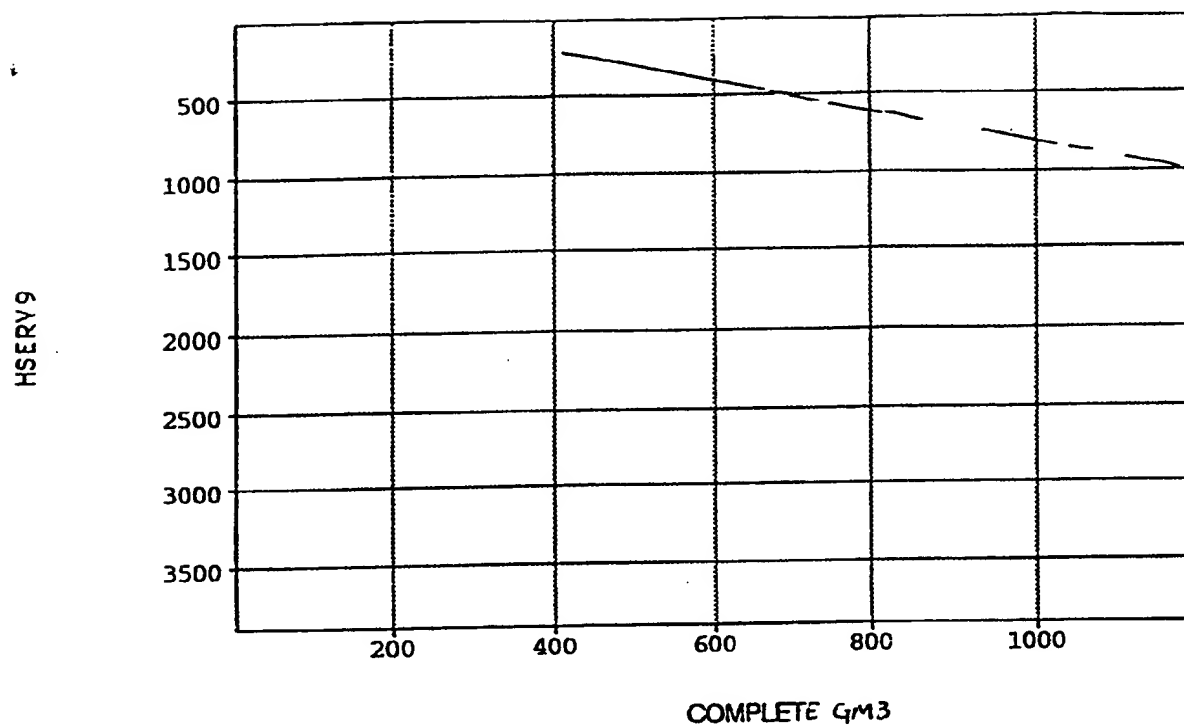
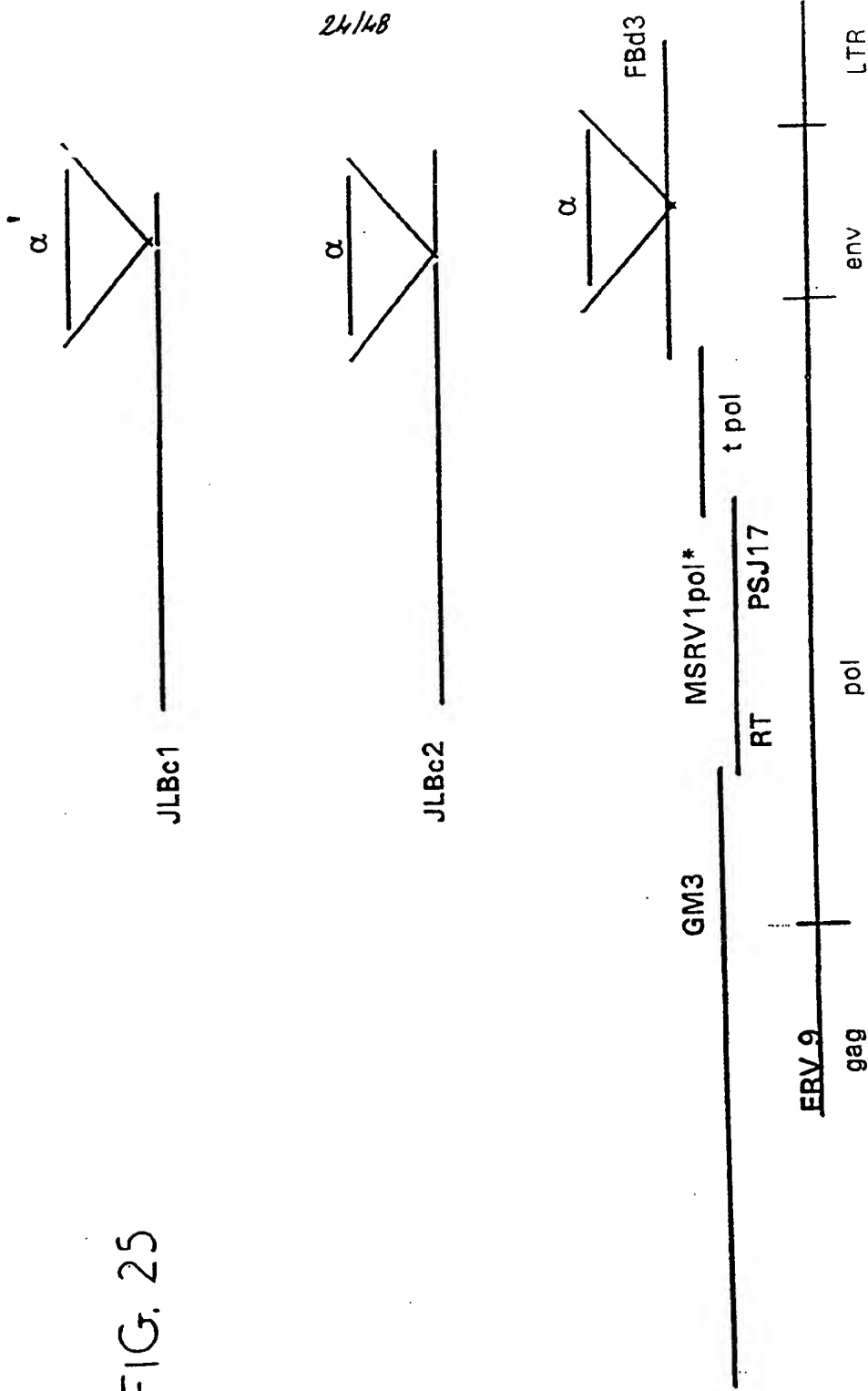


FIG. 24

FIG. 25



25/48

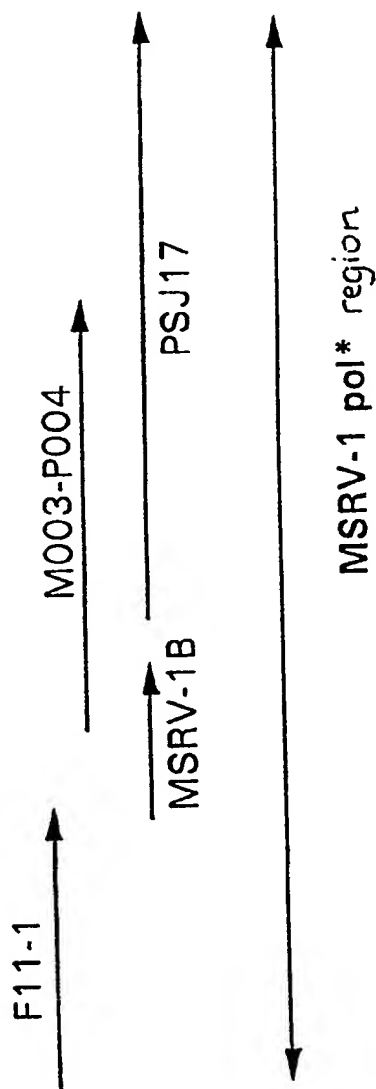


FIG. 26

26/48

FIG. 27a

SEQ ID NO 57 (POL)

90  
 ATG ATC CAG CAG CAC NGA GGG TGC CCG GGG CAA GCG CCA GGC CAT GCC ATC ACC CTC ACA GAG CCC CAG GTA TGC TTG ACC ATT GAG  
 M I Q Q Q D X G C P G Q A P A H A I T L T E P Q V C L T T I E  
 180  
 GGT CAG AAG GGT NAC TGT CTC CTG GAC ACT GGC GGN GGC TTC TCA GTC TTA CTT TCC TGT CCT GGA CAA CTG TCC TCC AGA TCT GTC ACT  
 G Q K G X C L L D T G G A F S V L L S C P G Q L S S R S V T  
 270  
 GTC CGA GGG GTC CTA GGA CAG CCA GTC ACT AGA TAC TTC TCC CAG CCA CTA AGT TGT GAC TCG GGA ACT TTA CTC TTC CCA CAT GCT TTT  
 V R G V L G Q P V T R Y F S Q P L S C D W G T L L F P H A F  
 360  
 CTA ATT ATG CCT GAA AGC CCC ACT CTC TTG TTG GGG AGA GAC ATT CTA GCA AAA GCA GGG GCC ATT ATA CAT GTG AAT ATA GGA GAA CGA  
 L I M P E S P T L L L G R D I L A K A G A I I H V N I G E G  
 450  
 ACA ACT GTT TGT CCC CTG CTT GAG GAA CGA ATT ANT CCT GAA GTC CCG GCA ACA GAA CGA TAT GGA CAA GCA AAG AAT GCC CGT  
 T T V C C P L L E E G I N P E V R A T E G Q Y G Q A K N A R  
 540  
 CCT GTT CAA GTT AAA CTA AAG GAT TCC ACC TCC TTT CCC TAC CAA AGG CAG TAC CCC CTC AGA CCC GAG ACC CAA CAA GAA CTC CAA AAG  
 P V Q V K L K D S T S F P Y Q R Q Y P L R P E T Q Q E L Q K  
 630  
 ATT GTA AAG GAC CTA AAA GGC CTA GTA AAA CCA AGC ANT AGC CCT TGC AAG ACT CCA ATT TTA GGA GTA AGG AAA CCC AAC CGA  
 I V K D L K A Q G L V K P S N S P C K T P I L G V R K P N G  
 720  
 CAG TCG AGG TTA GTG CAA GAA CTC AGG ATT ATC AAT GAG GCT GTT GTT CCT CTA TAC CCA CCT GTA CCT AAC CCT TAT ACA GTG CTT TCC  
 Q W R L V Q E L R I I N E A V V P L Y P A V P N P Y T V L S  
 810  
 CAA ATA CCA GAG GAA GCA GAG TGG TTT ACA GTC CTG GAC CTT AAG GAT GCC TTT TTC TGC ATC CCT GTA CCT P V R P D S Q F L  
 Q I P E E A E W F T V L D L K D A F F C I P V R P D S Q F L  
 900  
 TTT GCC TTT GAA GAT CCT TTG AAC CCA ACG TCT CAA CTC ACC TCG ACT GTT TTA CCC CAA GGG TTC AGG GAT AGC CCC CAT CTA TTT GGC  
 F A F E D P L N P T S Q L T W T V L P Q G F R D S P H L F G  
 990  
 CAG GCA TTA GCC CAA GAC TTG AGT CAA TTC TCA TAC CTG GAC ACT CTT GTC CTT CAG TAC ATG GAT GAT TTA CTT TTA GTC GCC CGT TCA  
 Q A L A Q D L S Q F S Y L D T L V L Q Y M D D L L L V A R S  
 1080  
 GAA ACC TTG TGC CAT CAA GCC ACC CAA GAA CTC TTA ACT TTC CTC ACT ACC TGT GGC TAC AAG GTT TCC AAA CCA AAG GCT CGG CTC TGC  
 E T L C H Q A T Q E L L T F L T C G Y K V S K P K A R L C

27/48

FIG. 27b

SEQ ID NO 57 (POL)

1170 TCA CAG GAG ATT AGA TAC TNA GGG CTA AAA TTA TCC AAA GGC ACC AGG GCC CTC AGT GAG GAA CGT ATC CAG CCT ATA CTG GCT TAT CCT  
 S Q E I R Y X G L K L S K G T R A L S E E R I Q P I L A Y P  
 1260 CAT CCC AAA ACC CTA AAG CAA CTA AGA GGG TTC CTT GGC ATA ACA GGT TTC TGC CGA AAA CAG ATT CCC AGG TAC ASC CCA ATA GCC AGA  
 H P K T L K Q L R G F L G I T G F C R K Q I P R Y X P I A R  
 1350 CCA TTA TAT ACA CTA ATT ANG GAA ACT CAG AAA GCC AAT ACC TAT TTA GTA AGA TGG ACA CCT ACA GAA GTG GCT TTC CAG GCC CTA AAG  
 P L Y T L I X E T Q K A N T Y L V R W T P T E V A F Q A L K  
 1440 AAG GCC CTA ACC CAA GCC CTA GTC TTC AGC TTG CCA ACA GGG CAA GAT TTT TCT TTA TAT GCC ACA GAA AAA ACA GGA ATA GCT CTA GGA  
 K A L T Q A P V F S L P T G Q D F S L Y A T E K T G I A L G  
 1530 GTC CTT ACG CAG GTC TCA GGG ATG AGC TTG CAA CCC GTG GTA TAC CTG AGT AAG GAA ATT GAT GTA GTG CCA AAG GGT TGG CCT CAT NGT  
 V L T Q V S G M S L Q P V Y L S K E I D V V A K G W P H X  
 1620 TTA TGG GTA ATG GAG GCA GTA GTC CCA GTC TNA GTA TCT GAA CCA GTT AAA ATA ATA CAG GGA AGA GAT CTT NCT GTG TGG ACA TCT CAT GAT  
 L W V M X A V A V X V S E A V K I I Q G R D L X V W T S H D  
 1710 GTG AAC GCC ATA CTC ACT GCT AAA GGA GAC TTG TGG TTG TCA GAC AAC CAT TTA CTT AAN TAT CAG GCT CTA TTA CTT GAA GAG CCA GTG  
 V N G I L T A K G D L W L S D N H L L X Y Q A L L L E E P V  
 1800 CTG NGA CTG CGC ACT TGT CCA ACT CTT AAA CCC AAA CTT ATG CTG CCC AGA AGC ATC TTT NTA CAG GTC CCC TTA GCC AAC CCT GAC CTC  
 L X L R T C A T L K P K L M L P R R I F X E V P L A N P D L  
 1890 AAC TAT ATA TAT ACT GAT CGA AGT TCG TTT GTA GAA AAG GGA TTA CAA AGG GNA GGA TAT NCC ATA GGT GTT AGT GAT AAA GCA GTA CTT  
 N Y I Y T D G S S F V E K G L Q R X G Y X I G V S D K A V L  
 1980 GAA AGT AAG CCT CTT CCC CAG GGA CCA CCG CCC TTA GCA GAA CTA GTG CCA CTG ACC CCG CGA GCT TTA GAA CTT TGG AAA GGG  
 E S K P L P Q G P A P P L A E L V A L T P R A L E L W K G  
 2070 AGG AGG ATA AAT GTG TAT ACA GAT AGC AAG TAT CCT TAT CTA ATC CGA AAT GCC CAT GGT GCA  
 R I N V Y T D S K Y A Y L I R N A H V V Y L I R N A H V A  
 2160 ATA TCG AAA GAA ACG CAG TTC CTA ACC TCT GGG GGA ACC CTT AAA TAC CAC AAG TTA ATC ATG GAG TTA TTG CAC ACA GTG CAA AAA  
 I W K E R E F L T S G G T P I K Y H K L I M E L L H T V Q K



28/48

SEQ ID NO 57 (POL)

FIG. 27c

CTC AAG GAG GTG GAA GTC TTA CAC TCC CAA AGC CAC CAA AGG CAA AGG CCA AGC CAA CCA AAG  
L K E V E V L H C Q S H Q K R E R G E E Q H K K W L Q R Q G K 2250

ACT AGC AGA AAG GAA AGA GAG AAA GAG ACA GAA AGT CAG AGA GAG CAA GAG ACA GAG CAC AAA GAG GCG GTC AGA GAG AGA GAG  
T S R K E R E K E T E S Q R E R E E E T E H K E G V R E R E 2340

AGA CAG AGA GTC AGA GAG AAG GAA AGA GAG AGA CCA AGA GAG CAA AAA GAA TGA  
R Q R V R E K E R E R G R D K E 2391

29/48

FIG. 28

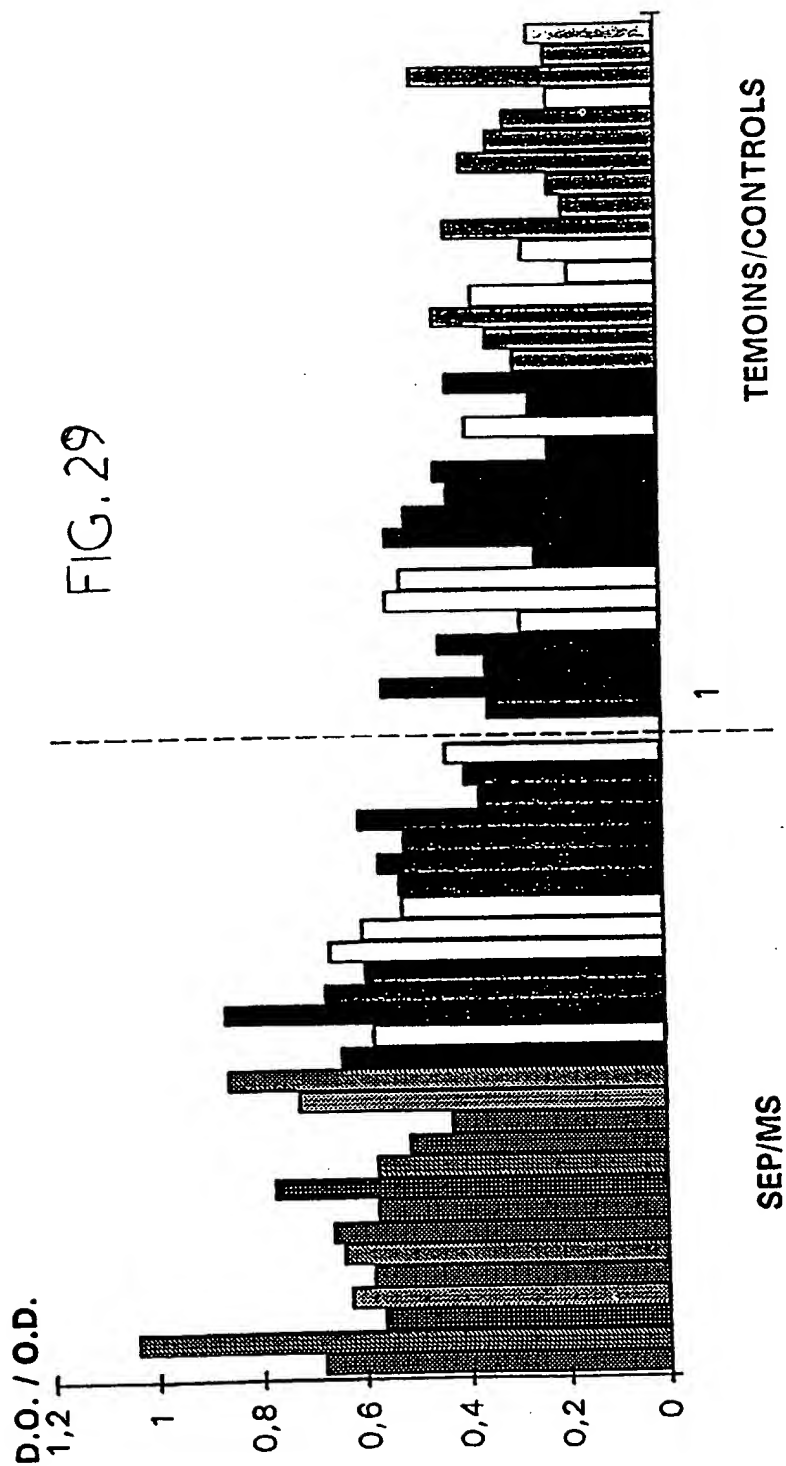
GATGCCTTTTTCTGCATCCCTGTACGTCCTGACTCTCAATTCTTGTTTGCCTTTGAAG  
ATCCTTTGAACCCAACGTCTCAACTCACCTGGACTGTTTTACCCCAAGGGTTCAGGGA  
TAGCCCCATCTATTTGGCCAGGCATTAGCCCAAGATGCCTTTTGCATCCCTGTACGTG  
ACTCTCAATTCTTGTTTGCCTTTGCCTTTGAAGATGCTTTGAACCCAACGTCTCAACT  
CACCTGGACTGTTTTACGCCAAGGGTTCAGGGATAGCCCCATCTATTTGGC  
CAGGCATTAGCCCAA

SEQ ID NO 40

Asp-Ala-Phe-Phe-Cys-Ile-Pro-Val-Arg-Pro-Asp-Ser-Gln-Phe-  
Leu-Phe-Ala-Phe-Glu-Asp-Pro-Leu-Asn-Pro-Thr-Ser-Gln-Leu-  
Thr-Trp-Thr-Val-Leu-Pro-Gln-Gly-Phe-Arg-Asp-Ser-Pro-His-  
Leu-Phe-Gly-Gln-Ala-Leu-Ala-Gln

SEQ ID NO 39 (POL2B)

30/48



31/48

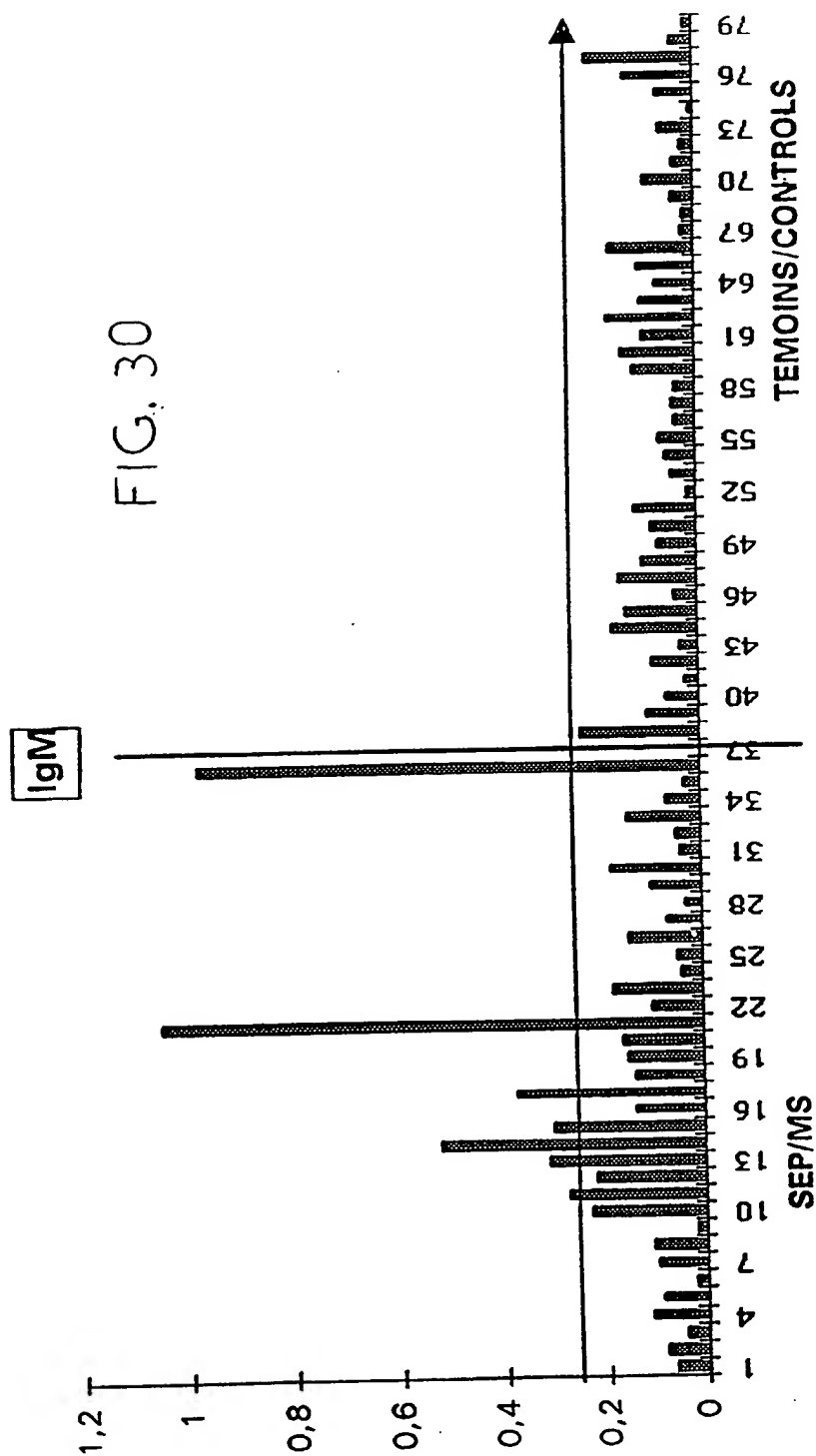
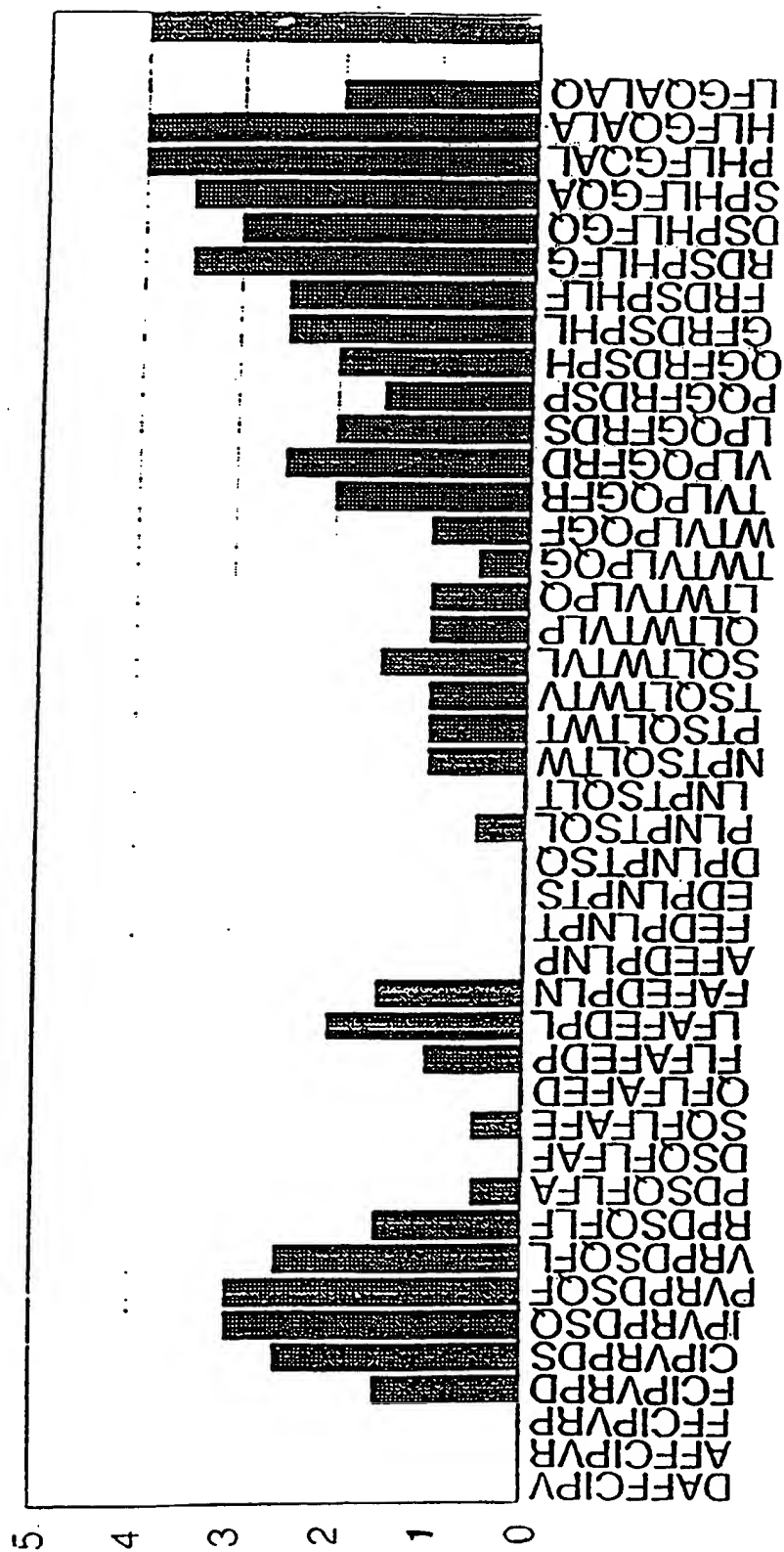
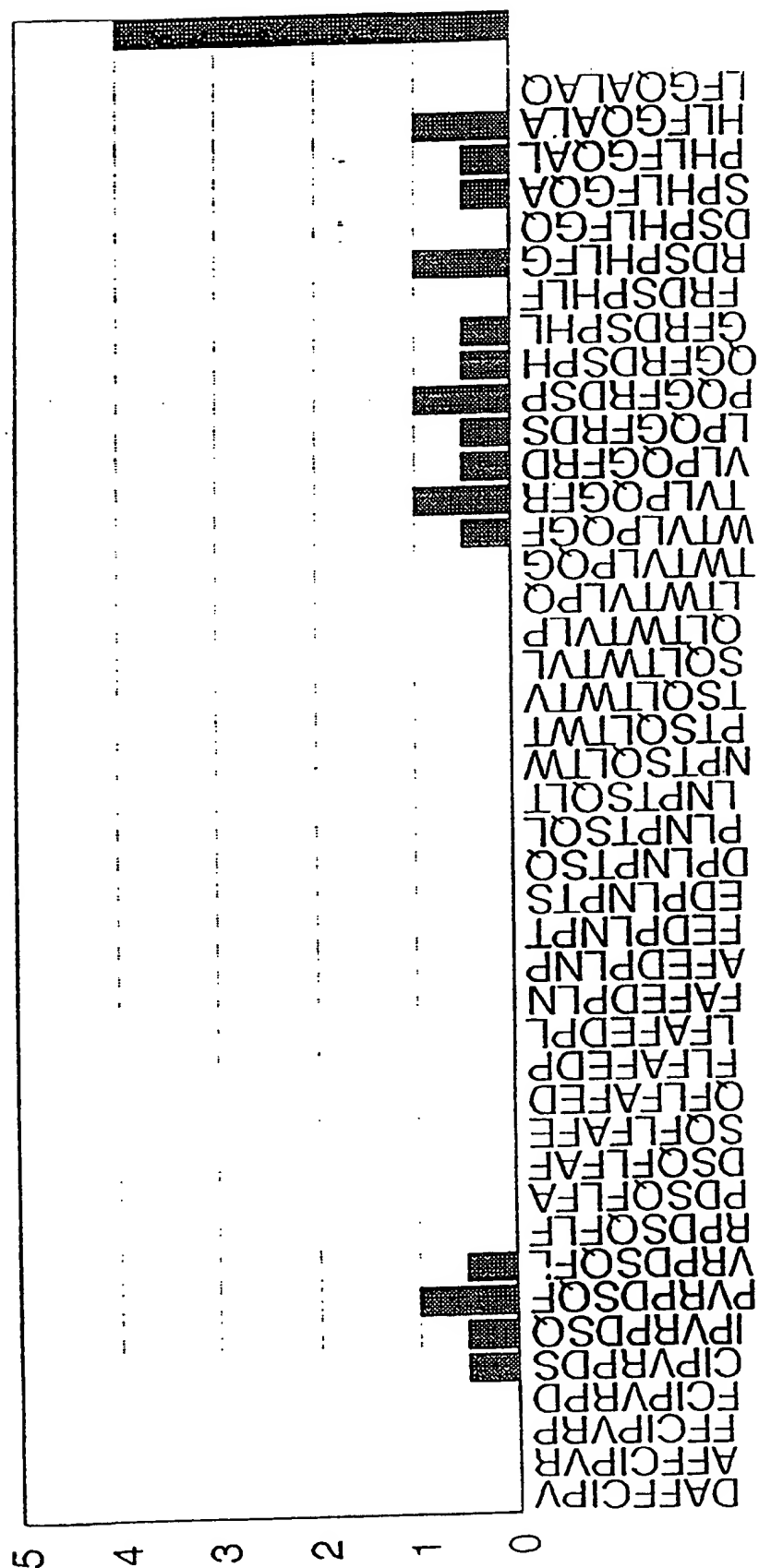


FIG. 31



33/48

FIG. 32





*35/48*

FIG. 34

Cys-Ile-Pro-Val-Arg-Pro-Asp-Ser-Gln-Phe-Leu SEQ ID NO 41

Val-Leu-Pro-Gln-Gly-Phe-Arg-Asp-Ser-Pro-His-Leu-Phe-Gly-  
Gln-Ala-Leu-Ala SEQ ID NO 42

Leu-Phe-Ala-Phe-Glu-Asp-Pro-Leu SEQ ID NO 43

Phe-Ala-Phe-Glu-Asp-Pro-Leu-Asn SEQ ID NO 44



FIG 35

36/48

	10	20	30	40	50	
	1234567890	1234567890	1234567890	1234567890	1234567890	
	CTTCCCAAC	TAATAAGGAC	CCCCCTTCA	ACCCAAACAG	TCCAAAAGGA	50
	L P Q L	I R T	P L S	T Q T V	Q K D	
	F P N	. . G P	P F Q	P K Q	S K R T	
	S P T	N K D	P P F N	P N S	P K G	
	CATAGACAAA	GGAGTAAACA	ATGAACCAAA	GAGTGCCAAT	ATTCCCTGGT	100
	I D K	G V N N	E P K	S A N	I P W L	
	. T K E	. T M N	Q R	V P I	F P G	
	H R Q R	S K Q	. T K	E C Q Y	S L V	
	TATGCAOCC	TCCAAGCGGTG	GGAGAAGAAT	TGGGCCCAGC	CAGAGTGCAT	150
	C T L	Q A V	G E E F	G P A	R V H	
	Y A P S	K R W	E K N	S A Q P	E C M	
	M H P	P S G G	R R I	R P S	Q S A C	
	GTACCTTTTT	CCTCTCTACA	CTTGAAGCAA	ATTAAATAG	ACNTAGGINA	200
	V P F S	L S H	L K Q	I K I D	X G X	
	Y L F	L S H T	. S K	L K .	T . V N	
	T F F	S L T	L E A N	. N R	X R X	
	ATTNTCAGAT	AGCCCTGATG	GYTATATIGA	TGTTTTACAA	GGATTAGGAC	250
	X S D	S P D G	Y I D	V L Q	G L G Q	
	X Q I	A L M	X I L M	F Y K	D . D	
	I X R .	P . W	L Y .	C F T R	I R T	
	AATCCTTTGA	TCTGACATGG	AGAGATATAA	TATTACTGCT	AAATCAGAGC	300
	S F D	L T W	R D I I	L L L	N Q T	
	N P L I	. H G	E I .	Y Y C .	I R R	
	I L .	S D M E	R Y N	I T A	K S D A	
	CTAACCTCAA	ATGAGAGAAG	TGCTGOCATA	ACTGGAGGCC	GAGAGTTTGG	350
	L T S N	E R S	A A I	T G A R	E F G	
	. P Q	M R E V	L P .	L E P	E S L A	
	N L K	. E K	C C H N	W S P	R V W	
	CAATCTCTGG	TATCTCAGTC	AGGTCAATGA	TAGGATGACA	ACGGAGGAAA	400
	N L W	Y L S Q	V N D	R M T	T E E R	
	I S G	I S V	R S M I	G . Q	R R K	
	Q S L V	S Q S	G Q .	. D D N	G G K	
	GAGAACGATT	CCCCACAGGG	CAGCAGGCAG	TTCOCAGTGT	AGCTCTCAT	450
	E R F	P T G	Q Q A V	P S V	A P H	
	E N D S	P Q G	S R Q	F P V .	L L I	
	R T I	P H R A	A G S	S Q C	S S S L	
	TGGACACAG	AATCAGAACA	TGGAGATTGG	TGCGCAGAC	ATTTA	495
	W D T E	S E H	G D W	C R R H	L	
	G T Q	N Q N M	E I G	A A D	I	
	G H R	I R T	W R L V	P Q T	F	

FIG 36

37/48

	10	20	30	40	50	
	1234567890	1234567890	1234567890	1234567890	1234567890	
	CTTCCCCAAC	TAATAAGGAC	CCCCCTTTCA	ACCCAAACAG	TOCAAAAGGA	50
	L P Q L	I R T	P L S	T Q T V	Q K D	
	CATAGACAAA	GGAGTAAACA	ATGAACCAAA	GAGTGCCAAT	ATTCCCTGGT	100
	I D K	G V N N	E P K	S A N	I P W L	
	TATGCACCOCT	OCAAGCGGTG	GGAGAAGAAT	TCGGCCCCAGC	CAGAGTGCAT	150
	C T L	Q A V	G E E F	G P A	R V H	
	GTACCTTTT	CTCTCTCACA	CTGAAGCAA	ATTAAAATAG	ACCTAGGTAA	200
	V P F S	L S H	L K Q	I K I D	L G K	
	ATTCTCAGAT	AGCCCTGATG	GYTATATTGA	TGTTTACAA	GGATTAGGAC	250
	F S D	S P D G	Y I D	V L Q	G L G Q	
	AATCCTTTGA	TCTGACATGG	AGAGATATAA	TATTACTGCT	AAATCAGAGG	300
	S F D	L T W	R D I I	L L L	N Q T	
	CTAACCTCAA	ATGAGAGAAG	TGCTGOCATA	ACTGGAGGCC	GAGAGTTTGG	350
	L T S N	E R S	A A I	T G A R	E F G	
	CAATCTCTGG	TATCTCAGTC	AGGTCAATGA	TAGGATGACA	ACGGAGGAAA	400
	N L W	Y L S Q	V N D	R M T	T E E R	
	GAGAAGGATT	CCCCACAGGG	CAGCAGGCAG	TTCCAGTGT	AGCTCTCAT	450
	E R F	P T G	Q Q A V	P S V	A P H	
	TGGACACAG	AATCAGAACA	TGGAGATTGG	TGCSCAGAC	ATTTACAAT	500
	W D T E	S E H	G D W	C R R H	L Q L	
	TGCGTCTAN	AAGGACTNAG	GAAAACTAGG	ARGACTANGA	ATTATTCAN	550
	A C X	K D X G	K L G	R L X	I I Q X	
	GATGTCCACT	ANNACACAGG	GGAAAGGAG	AAAATCTAC	TGCTTTCTG	600
	C P L	X H R	G K E E	N P T	A F L	
	GAGGACTAA	GGGAGGCATT	GAGGAAGCAT	ACCAGGCAAG	TGGACATTGG	650
	E R L R	E A L	R K H	T R Q V	D I G	
	AGGCTCTGGA	AAAGGGAAAA	GTGGGGGAAA	TTATATGCT	AATAGGGCTT	700
	G S G	K G K S	W A N	Y M P	N R A C	
	GCTTCCAGTG	CAGICTACAA	GGAAGCTTTA	GAAAAGATTG	TOCAAGTAGA	750
	F Q C	S L Q	G R F R	K D C	P S R	
	AATAAGCGGC	CCCTGTGTOCA	TGCCCCCTAT	GTCAAGGGAA	TCACTGGAG	800
	N K P P	L V H	A P Y	V K G I	T G R	
	GCTTACTGOC	CCAGGGGAGG	AAGGTCTCT	GAGTCAGAAG	CCACTTAACT	850
	P T A	P G D E	G P L	S Q K	P L T	
GA						852

38/48

FIG 37

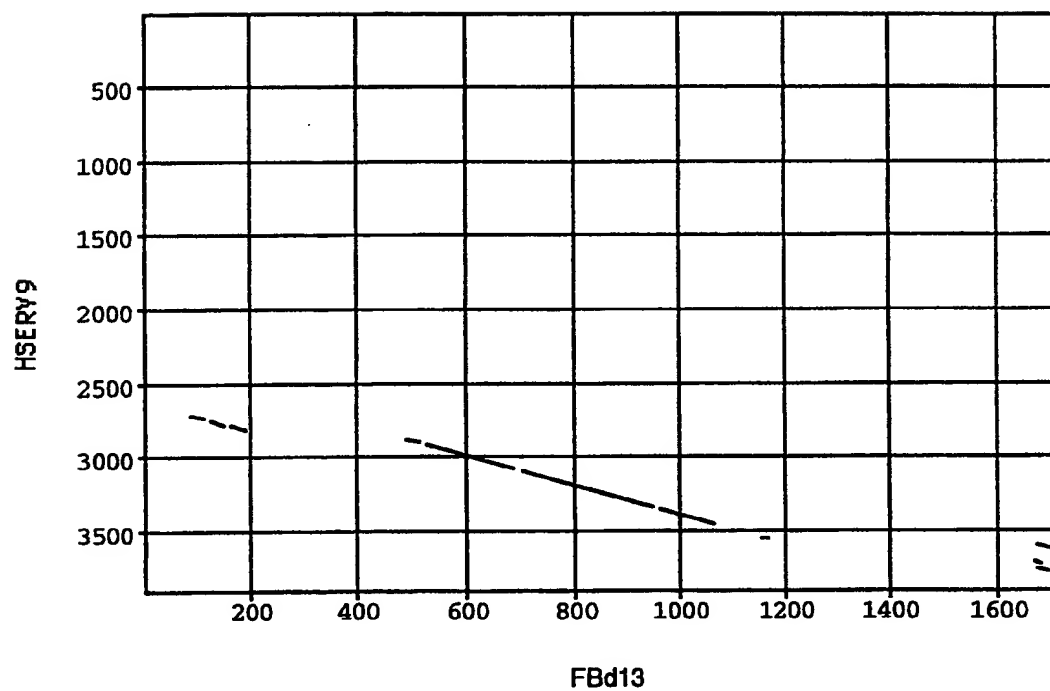


FIG 38  
a

39/48

	10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	
AAGGAACTC	AGAAAGCCAA	TACCCATTTA	GTAAAGATGGA	CACCAGAAGC		50
K E T Q	K A N	T H L	V R W T	P E A		
R K L	R K P I	P I .	. D G	H Q K Q		
G N S	E S Q	Y P F S	K M D	T R S		
AGAAGCAGCT	TTCCAGGCC	TAAAGAATC	OCTAACCCAA	GCCCCAGTGT		100
E A A	F Q A L	K K S	L T Q	A P V L		
K Q L	S R P	. R N P	. P K	P Q C		
R S S	F P G P	K E I	P N P S	P S V		
TAAGCTTGCC	AACGGGGCAA	GACTTTTCTT	TATATGTCAC	AGAAAAACAG		150
S L P	T G Q	D F S L	Y V T	E K Q		
. A C Q	R G K	T F L	Y M S Q	K N R		
K L A	N G A R	L F F	I C H	R K T G		
GAATAGCTCT	AGGAGTCTT	ACACAGGTCC	AAGGGACAAG	CTTGCAACCT		200
E . L .	E S L	H R S	K G Q A	C N L		
N S S	R S P Y	T G P	R D K	L A T C		
I A L	G V L	T Q V Q	G T S	L Q P		
GTTGCATACC	TGAGTAAGGA	AACIGATGTA	NIGGCAAAGG	GTGGCCCTCA		250
W H T	. V R K	L M X	W Q R	V G L I		
G I P	E . G N .	C X	G K G	L A S		
V A Y L	S K E	T D V	X A K G	W P H		
TGTGTTACAG	GTAGGGCAGC	AGTAGCAGTC	TTAGTTTCTG	AAACAGTTAA		300
V Y R	. G S	S S S L	S F .	N S .		
L F T G	R A A	V A V	L V S E	T V K		
C L Q	V G Q Q	. Q S	. F L	K Q L K		
AATAATACAG	GGAAGAGATC	TTACTGIGTG	GACATCTCAT	GATGIGAACG		350
N N T G	K R S	Y C V	D I S .	C E R		
I I Q	G R D L	T V W	T S H	D V N G		
. Y R	E E I	L L C G	H L M	M . T		
GCATACTCAC	TGCTAAAGAG	GACTGIGGGC	TGICAGACAA	CCATTACTT		400
H T H C	. R G	L V A	V R Q	P F T .		
I L T	A K E	D L W L	S D N	H L L		
A Y S L	L K R	T C G	C Q T T	I Y L		
AAATAGCAGG	TTCTATTACT	TGAAGTGGCA	GTGCTGGGAC	TGCACATTIG		450
I A G	S I T .	S A S	A A T	A H L		
K . Q	V L L L	E V P	V L R L	H I C		
N S R	F Y Y L	K C Q	C C D	C T F V		
TGCAACTCTT	AAACCAGCCA	CATTCTCTCC	AGACAATGAA	GAAAAGATAG		500
C N S .	P S H	I S S	R Q .	R K D R		
A T L	N P A T	F L P	D N E	E K I E		
Q L L	T Q P	H F F Q	T M K	K R .		

40/48

FIG38  
b

	10	20	30	40	50	
	1234567890	1234567890	1234567890	1234567890	1234567890	
	AACATAACTG	TCAACAAGTA	ATTGCTCAAA	CTATGCTGC	TGAGGGGAC	550
	T . L	S T S N	C S N	L C C	S R G P	
	H N C	Q Q V	I A Q T	Y A A	R G D	
	N I T V	N K .	L L K	P M L L	E G T	
	CTCTAGAGG	TTCCCTGAC	TGATCCCGAC	CTCAACTTGT	ATACTGATGG	600
	S R G	S L D .	S R P	Q L V	Y . W	
	L L E V	P L T	D P D	L N L Y	T D G	
	F . R	F P .	L I P T	S T C	I L M E	
	AAGTTCCTTG	GCAGAAAAAG	GACTTTGAAA	AGGGGGTAT	GCAGTGATCA	650
	K F L G	R K R	T L K	S G V C	S D Q	
	S S L	A E K G	L . K	A G Y	A V I S	
	V P W	Q K K	D F E K	R G M	Q . S	
	GTCATAATGG	AATACTTGAA	AGTAATCGCC	TCACTCCAGG	AACTAGTGCT	700
	. . W	N T . K	. S P	H S R	N . C S	
	D N G	I L E	S N R L	T P G	T S A	
	V I M E	Y L K	V I A	S L Q E	L V L	
	CACTGGCAG	AACTAATAGC	CTCACTTGG	GCACTAGAAT	TAGCAGAAGG	750
	P G R	T N S	P H L G	T R I	R R R	
	H L A E	L I A	L T W	A L E L	G E G	
	T W Q	N . .	P S L G	H . N	. E K E	
	AAAAAGGGTA	AATATATATT	CAGACTICAA	GTATGCTTAC	CTAGTCTCC	800
	K K G K	Y I F	R L .	V C L P	S P P	
	K R V	N I Y S	D S K	Y A Y	L V L H	
	K G .	I Y I	Q T L S	M L T	. S S	
	ATGCCCATGC	AGCAATATGG	AGAGAGAGGG	AATTCCTAAC	TTCTGAGGGA	850
	C P C	S N M E	R E G	I P N	F . G N	
	A H A	A I W	R E R E	F L T	S E G	
	M P M Q	Q Y G	E R G	N S .	L L R E	
	ACACCTATCA	ACCATCAGGG	AAGCATTAG	GAGATTATTA	TGGCTGTAC	900
	T Y Q	P S G	K P L G	D Y Y	W L Y	
	T P I N	H Q G	S H .	E I I I	G C T	
	H L S	T I R E	A I R	R L L	L A V Q	
	AGAAACCTAA	AGAGGIGGCA	GTCTTACACT	GCCAGGGTCA	TCAGGAAGAA	950
	R N L K	R W Q	S Y T	A R V I	R K K	
	E T .	R G G S	L T L	P G S	S G R R	
	K P K	E V A	V L H C	Q G H	Q E E	
	GAGGAAAGGG	AATAGAAGG	CAATCGCCAA	GCGGATATTG	AAGCAAAAAA	1000
	R K G	K . K A	I A K	R I L	K Q K K	
	G K G	N R R	Q S P S	G Y .	S K K	
	E E R E	I E G	N R Q	A D I E	A K K	

41/48

FIG 38  
C

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
AGCCGCAAGG	CAGGACTCTC	CATTAGAAAT	GCTTATAGAA	GGACCCCTAG	1050
P Q G R T L H .	K C L . K D P .				
S R K A G L S I R N	A Y R R T P S				
A A R Q D S P L E M	L I E G P L V				
TATGGGGTAA	TCCCTCTGG	GAAACCAAGC	CCCAGTACTC	AGCAGGAAAA	1100
Y G V I P S G K P S	P S T Q Q E K				
M G . S P L G N Q A	P V L S R K N				
W G N P L W E T K P	Q Y S A G K				
ATAGAATAGG	AAACCTCACA	AGGACATACT	TTCCTCCCCCT	CCAGATGGCT	1150
. N R K P H K D I L	S S P P D G .				
R I G N L T R T Y F	P P L Q M A				
I E . E T S Q G H T	F L P S R W L				
AGCCACTGAG	GAAGGAA				1167
P L R K E					
S H . G R					
A T E E G					

42/48

FIG 39

	10	20	30	40	50	
	1234567890	1234567890	1234567890	1234567890	1234567890	
	AACTTGGGIG CTAGAAGGAC TAAGGAAAAC TAGGAAGACT ATGAATTATT					50
	N L R A R R T K E N . E D Y E L F					
	T C V L E G L R K T R K T M N Y S					
	L A C . K D . G K L G R L . I I					
Q	CAATGATGTC CACTATAACA CAGGGGAAAG GAAGAAATC CTAAGTGCIT					100
	N D V H Y N T G E R K K I L L P F					
	M M S T I T Q G K G R K S Y C L					
	Q . C P L . H R G K E E N P T A F					
	TCTGGAGAGA CTAAGGGAGG CATTCAGGAA GCATAACAGG CAAGTGGACA					150
	W R D . G R H . G S I P G K W T					
	S G E T K G G I E E A Y Q A S G H					
	L E R L R E A L R K H T R Q V D I					
	TTGGAGGCTC TGGAAAAGGG AAAAGTTGGG CAAATTGAAT GCTAATAGG					200
	L E A L E K G K V G Q I E C L I G					
	W R L W K R E K L G K L N A . . G					
	G G S G K G K S W A N . M P N R					
	GCTTGCTTCC AGTGCAGTCT ACAAGGACGC TTATAGAAAAG ATTGTCCAAG					250
	L A S S A V Y K D A L E K I V Q V					
	L L P V Q S T R T L . K R L S K					
	A C F Q C S L Q G R F R K D C P S					
	TAGAAATAAG CCGCCCTCG TCATGCCCC TTATGICAAG GGAATCACTG					300
	E I S R P S S M P L M S R E S L					
	. K . A A P R P C P L C Q G N H W					
	R N K P P L V H A P Y V K G I T G					
	GAAGGCTAC TGGCCAGGG GACGAAGGIC CTCTGAGICA GAAGCCACTA					350
	E G L L P Q G T K V L . V R S H .					
	K A Y C P R G R R S S E S E A T N					
	R P T A P G D E G P L S Q K P L					
	ACCTGATCAT CCAGCAGCAG GACTGAGGGT GCGCGGGGCA AGTGCCAGCC					400
	P D D P A A G L R V P G A S A S P					
	L M I Q Q Q D . G C P G Q V P A					
	T . . S S S R T E G A R G K C Q P					
	CATGCCATCA CCTCAGAGC CCGGGGTAIG TTGACCATT GAGAGCCAGG					450
	C H H P Q S P G Y V . P L R A R					
	H A I T L R A P G M F D H . E P G					
	M P S P S E P R V C L T I E S Q E					
	AAGTTAACTG TCTCTGGAC ACTGGGCGAG CCTTCTCAGT CTACCTTCC					500
	K L T V S W T L A Q P S Q S Y F P					
	S . L S P G H W R S L L S L T F L					
	V N C L L D T G A A F S V L L S					

43/48

FIG 39  
b

	10	20	30	40	50	
	1234567890	1234567890	1234567890	1234567890	1234567890	
	TGTCOCAGAC	AATTGTCTC	CAGATCTGTC	ACTATCCGAG	GGGTCTAAG	550
	V P D	N C P P	D L S	L S E	G S . D	
	S Q T	I V L	Q I C H	Y P R	G P K	
	C P R Q	L S S	R S V	T I R G	V L R	
	ACAGCCAGTC	ACTACATACT	TCCTCAGCC	ACTAGTTGT	GACTGGGGAA	600
	S Q S	L H T	S L S H	. V V	T G E	
	T A S H	Y I L	L S A	T K L .	L G N	
	Q P V	T T Y F	S Q P	L S C	D W G T	
	CTTTACTCTT	TTCACATGCT	TTTCTAATTA	TGCTGAAAG	CCCCACTOCC	650
	L Y S F	H M L	F . L	C L K A	P L P	
	F T L	F T C F	S N Y A .	K P H S L		
	L L F	S H A	F L I M	P E S	P T P	
	TTGTTAGGGA	GAGACATTTT	AGCAAAAGCA	GGGGCCATTA	TACACTGAA	700
	C . G	E T F .	Q K Q	G P L	Y T . T	
	V R E	R H F	S K S R	G H Y	T P E	
	L L G R	D I L	A K A	G A I I	H L N	
	CATAGCAAAA	GGATACCCA	TTTGTCTGTC	CTGCTTGAG	GAAGCAATTA	750
	. E K	E Y P	F A V P	C L R	K E L	
	H R K R	N T H	L L S	P A .	G R N .	
	I G K	G I P I	C C P	L L E	E G I N	
	ATCTGAAGT	CTGGGCAATA	GAAGGACAT	ATGGACAAGC	AAAGATGCC	800
	I L K S	G Q .	K D N	M D K Q	R M P	
	S . S	L G N R	R T I	W T S	K E C P	
	P E V	W A I	E G Q Y	G Q A	K N A	
	CGTCTGTTT	AAGTTAACT	AAAGGATCT	GGTCTCTTC	CTACCAAG	850
	V L F	K L N .	R I L	P P F	P T K G	
	S C S	S . T	K G F C	L L S	L P K	
	R P V Q	V K L	K D S	A S F P	Y Q R	
	GAAGTACCT	CTTAGACCCG	AGGCCCTACA	AGGACTCAAA	AGATTGTTAA	900
	S T L	L D P	R P Y K	D S K	D C .	
	E V P S	. T R	G P T	R T Q K	I V K	
	K Y P	L R P E	A L Q	G L K	R L L R	
	GGACTTAAAA	GOCCAAGGOC	TAGTAAAAOC	ATGCAGTAGC	COCTGCAATA	950
	G P K S	P R P	S K T	M Q .	P L Q Y	
	D L K	A Q G L	V K P	C S S	P C N T	
	T . K	P K A .	. N H	A V A	P A I	
	CTCCATTTT	AGGAGTAAGG	AAACCCAAAG	GACAGTGGAG	GTTAGTGCAA	1000
	S N F	R S K E	T Q R	T V E	V S A R	
	P I L	G V R	K P N G	Q W R	L V Q	
	L Q F .	E . G	N P T	D S G G	. C K	



FIG 39  
C

44/48

	10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	
GATCTCAGGA	TTATTATGA	GGCTGTTTT	CCTCTATACC	CAGCTGTATC		1050
S Q D Y . .	G C F S	S I P S C I				
D L R I I N E	A V F	P L Y P A V S				
I S G L L M R	L F F	L Y T Q L Y L				
TAGCCCTAT	ACTCTGCTTT	CCCTAATACC	AGAGGAAGCA	GAGTAGTTA		1100
. P L Y S A F	P N T R G S R	V V Y .				
S P Y T L L S	L I P E E A E .	F T				
A L I L C F P .	Y Q R K Q S S L					
CAGTCTGGA	CCTTAAGGAT	GGCTCTTTCT	GCATCCCTGT	ACATCCIGAT		1150
S P G P . G C	L F L H P C T S .	F				
V L D L K D A	S F C I P V H P D					
Q S W T L R M	P L S A S L Y I L I					
TCTCAATTCT	TGTTTGCTT	TGAAGATCCT	TTGAACCCAA	TGTCCTAATT		1200
S I L V C L .	R S F E P N V S I					
S Q F L F V F	E D P L N P M S Q F					
L N S C L S L	K I L . T Q C L N S					
CACCTGGACT	GTTTTACCC	AGGGGTTCGG	GGATAGCCCC	CATCTATTTC		1250
H L D C F T P	G V P G . P P S I W					
T W T V L P Q	G F R D S P H L F G					
P G L F Y P	R G S G I A P I Y L					
GCCAGGCATT	AGCCCAAGAC	TTCAGCCATT	TCTCATACT	GGACATCTTC		1300
P G I S P R L	E P I L I P G H L V					
Q A L A Q D L	S Q F S Y L D I L					
A R H . P K T .	A N S H T W T S C					
TCCTTCGGTA	TGGGATGATT	TAATTTTAGC	CACCGTTCA	GAAACCTTGT		1350
L R Y G M I .	F . P P V Q K P C					
S F G M G .	F N F S H P F R N L V					
P S V W D D L	I L A T R S E T L C					
GCCATCAAGC	CACCCAAGCG	TTCCTAAATT	TCCTCACTCC	GIGTGGCTAC		1400
A I K P P K R	S . I S S L R V A T					
P S S H P S V	L K F P H S V W L Q					
H Q A T Q A	F L N F L T P C G Y					
AAGGTTTCCA	AACCAAGGC	TCAGCTCTGC	TCACAGCAGG	TAAATACTT		1450
R F P N Q R L	S S A H S R L N T .					
G F Q T K G	S A L L T A G . I L					
K V S K P K A	Q L C S Q Q V K Y L					
AGGGTTAAAA	TTATCCAAG	GCAACAGGC	CCTCTGTGAG	GAATGTATCC		1500
G . N Y P K	A P G P S V R N V S					
R V K I I Q R	H Q G P L . G M Y P					
G L K L S K G	T R A L C E E C I Q					

45/48

FIG 39  
d

10	20	30	40	50
1234567890	1234567890	1234567890	1234567890	1234567890
AAOCTGTACT	GGCTTATCTT	CATCCCAAAA	CCCTAAAGCA	ACTAAGAAGG
N L Y W	L I F	I P K	P . S N	. E G
T C T	G L S S	S Q N	P K A	T K K V
P V L	A Y L	H P K T	L K Q	L R R
TOCTTGGCAT AACAGGTTTC TGCCGAA				
P W H N R F L P				
L G I T G F C R				
S L A . Q V S A E				

1550

1577

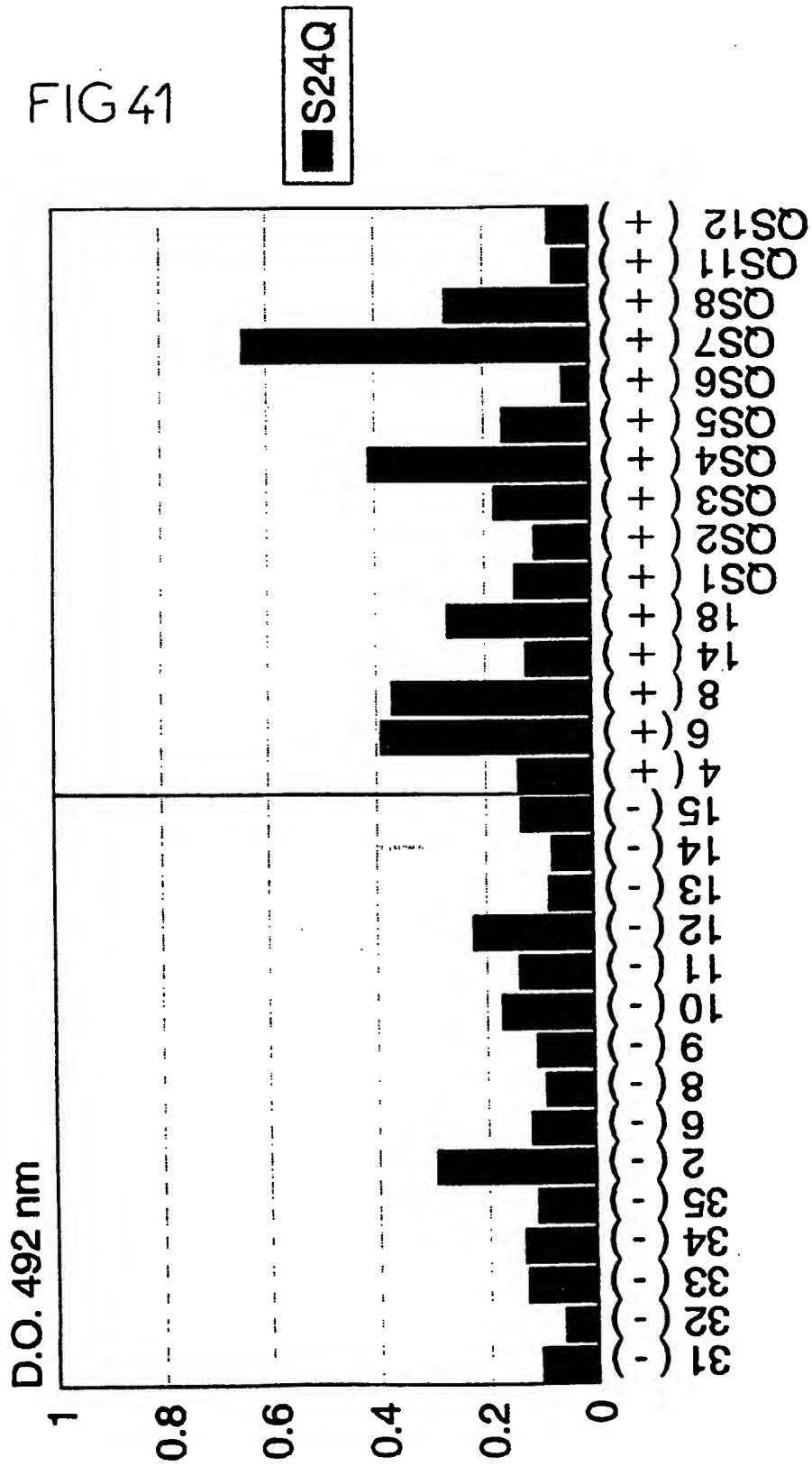
46/48

## FIG 40

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
TCACGACGA	GGACTGAGG	TGCCCCGGGC	AAGTGCCAGC	CCATGCCATC	50
S S S R	T E G	A R G	K C Q P	M P S	
ACCTCAGAG	CCCCGGGAT	GTTTGACCAT	TGAGAGCCAG	GAAGTTAACT	100
P S E	P R V C	L T I	E S Q	E V N C	
GTCCTCGGA	CACIGGGCA	GCCTTCACAG	TCTTACTTTC	CTGTCCAGA	150
L L D	T G A	A F S V	L L S	C P R	
CAATTGTCCT	CCAGATCTGT	CACATACCGA	GGGGTCTAA	GACAGCCAGT	200
Q L S S	R S V	T I R	G V L R	Q P V	
CACATACAT	TTCTCTCAGC	CACATAAGTG	TGACTGGGGA	ACTTTACTCT	250
T T Y	F S Q P	L S C	D W G	T L L F	
TTTCACATGC	TTTTCTAATT	ATGCCIGAAA	GCCCCACTCC	CTGTGTAGGG	300
S H A	F L I	M P E S	P T P	L L G	
AGAGACATTT	TAGCAAAAGC	AGGGGGCCATT	ATACACCTGA	ACATAGGAAA	350
R D I L	A K A	G A I	I H L N	I G K	
AGGAATACCC	ATTTCCTGTC	CCCTGCTTGA	GGAAGGAATT	AATCCTGAAG	400
G I P	I C C P	L L E	E G I	N P E V	
TCIGGGCAAT	AGAAGGACAA	TATGGACAAG	CAAAGAATGC	CGTCCCTGTT	450
W A I	E G Q	Y G Q A	K N A	R P V	
CAAGTTAAAC	TAAAGGATTC	TGCCCTCTTT	CCCTACCAAA	GGAAGTACCC	500
Q V K L	K D S	A S F	P Y Q R	K Y P	
TCTTAGACCC	GAGGCCCTAC	AAGGACTCAA	AAGATTGTTA	AGGAOCT	547
L R P	E A L Q	G L K	R L L	R T	

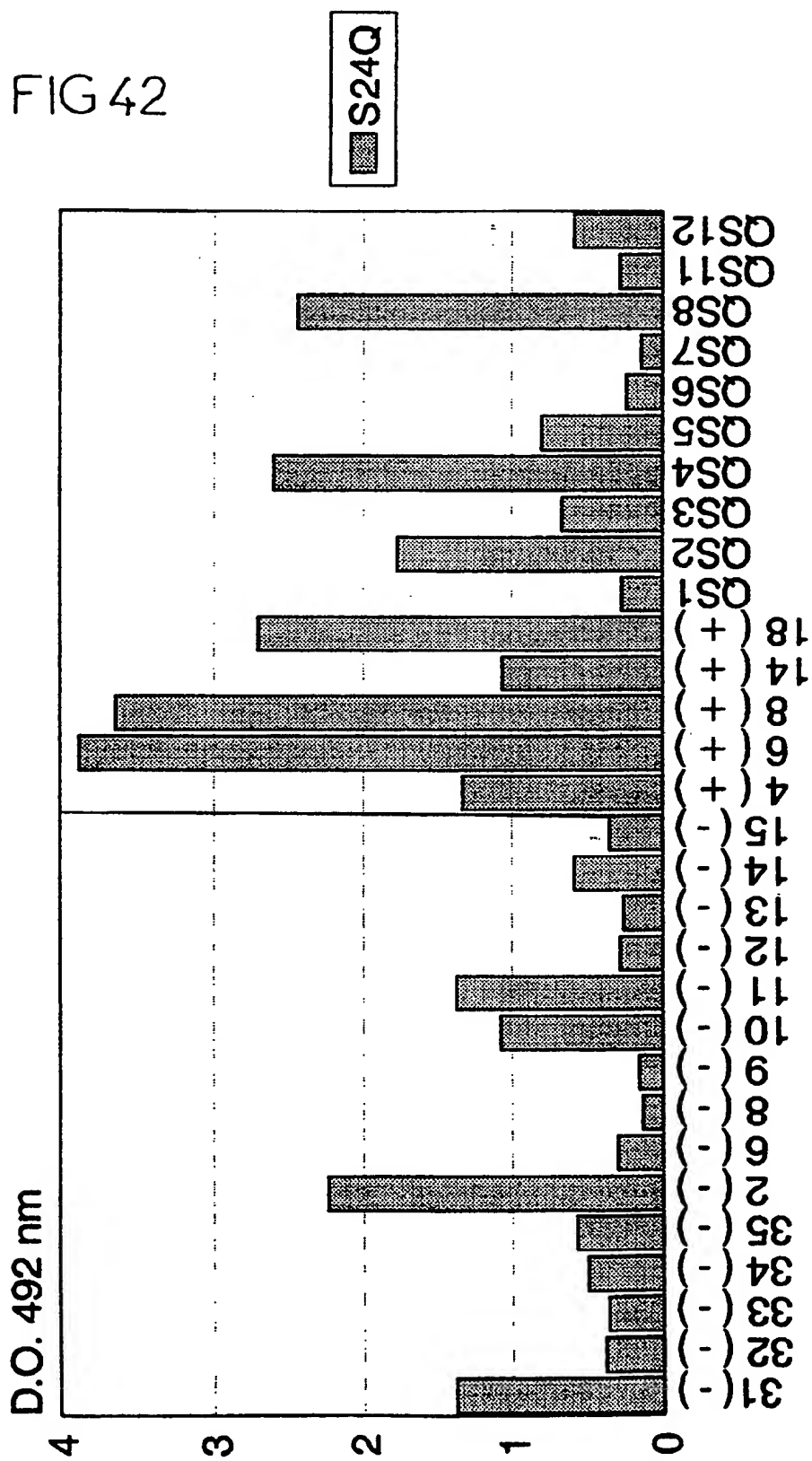
47/48

FIG 41



48/48

FIG 42



# INTERNATIONAL SEARCH REPORT

Int'l Application No  
PCT/FR 96/01244

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/48 C12N7/00 C07K14/15 C12Q1/68 C07K16/10  
G01N33/569 A61K39/21 A61K39/42 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C07K C12N C12Q G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>WO,A,94 28138 (UNIVERSITY COLLEGE LONDON) 8 December 1994 see page 1, line 25 - page 3, line 5 --- -/--</p>	1-46

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

20 November 1996

Date of mailing of the international search report

26. 11. 96

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+ 31-70) 340-3016

Authorized officer

Montero Lopez, B

## INTERNATIONAL SEARCH REPORT

Int onal Application No

PCI/FR 96/01244

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	RESEARCH IN VIROLOGY, vol. 143, no. 5, 1992, pages 337-350, XP000569296 H. PERRON ET AL.: "In vitro transmission and antigenicity of a retrovirus isolated from a multiple sclerosis patient " see page 338, left-hand column, paragraph 1 - paragraph 3 see page 342, right-hand column, paragraph 2 - page 345, right-hand column, paragraph 1 see page 346, right-hand column, paragraph 2 - page 347, left-hand column, paragraph 1 see page 348, left-hand column, paragraph 2 ---	25-35
A	LANCET THE, vol. 337, 6 April 1991, LONDON GB, pages 862-863, XP002001596 H. PERRON ET AL.: "Isolation of retrovirus from patients with multiple sclerosis" see page 862, left-hand column, paragraph 2 see page 862, right-hand column, paragraph 2 -----	1-19, 24-37, 39-46

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PC1/FR 96/01244

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9428138	08-12-94	AU-A- 6760094	20-12-94
		CA-A- 2163641	08-12-94
		EP-A- 0700441	13-03-96
-----			



# RAPPORT DE RECHERCHE INTERNATIONALE

Der: Internationale No  
PC1/FR 96/01244

<b>A. CLASSEMENT DE L'OBJET DE LA DEMANDE</b> CIB 6    C12N15/48    C12N7/00    C07K14/15    C12Q1/68    C07K16/10 G01N33/569    A61K39/21    A61K39/42    A61K48/00					
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB					
<b>B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE</b> Documentation minimale consultée (système de classification suivi des symboles de classement) CIB 6    C07K    C12N    C12Q    G01N    A61K					
Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche					
Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés)					
<b>C. DOCUMENTS CONSIDERES COMME PERTINENTS</b>					
Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents				no. des revendications visées
A	W0,A,94 28138 (UNIVERSITY COLLEGE LONDON) 8 Décembre 1994 voir page 1, ligne 25 - page 3, ligne 5 ---				1-46
A	RESEARCH IN VIROLOGY, vol. 143, no. 5, 1992, pages 337-350, XP000569296 H. PERRON ET AL.: "In vitro transmission and antigenicity of a retrovirus isolated from a multiple sclerosis patient " voir page 338, colonne de gauche, alinéa 1 - alinéa 3 voir page 342, colonne de droite, alinéa 2 - page 345, colonne de droite, alinéa 1 voir page 346, colonne de droite, alinéa 2 - page 347, colonne de gauche, alinéa 1 voir page 348, colonne de gauche, alinéa 2 --- <div style="text-align: right;">-/--</div>				25-35
<div style="display: flex; justify-content: space-between;"> <span><input checked="" type="checkbox"/> Voir la suite du cadre C pour la fin de la liste des documents</span> <span><input checked="" type="checkbox"/> Les documents de familles de brevets sont indiqués en annexe</span> </div>					
<div style="display: flex;"> <div style="flex: 1;"> <p>* Catégories spéciales de documents cités:</p> <p>"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent</p> <p>"E" document antérieur, mais publié à la date de dépôt international ou après cette date</p> <p>"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)</p> <p>"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens</p> <p>"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée</p> </div> <div style="flex: 1;"> <p>"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention</p> <p>"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément</p> <p>"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier</p> <p>"&amp;" document qui fait partie de la même famille de brevets</p> </div> </div>					
Date à laquelle la recherche internationale a été effectivement achevée  <div style="text-align: center; font-weight: bold;">20 Novembre 1996</div>			Date d'expédition du présent rapport de recherche internationale  <div style="text-align: center; font-weight: bold;">26. 11. 96</div>		
Nom et adresse postale de l'administration chargée de la recherche internationale: Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tél. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016			Fonctionnaire autorisé  <div style="text-align: center; font-weight: bold;">Montero Lopez, B</div>		

# RAPPORT DE RECHERCHE INTERNATIONALE

Den Internationale No  
PC1/FR 96/01244

C(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	<p>LANCET THE, vol. 337, 6 Avril 1991, LONDON GB, pages 862-863, XP002001596 H. PERRON ET AL.: "Isolation of retrovirus from patients with multiple sclerosis" voir page 862, colonne de gauche, alinéa 2 voir page 862, colonne de droite, alinéa 2 -----</p>	<p>1-19, 24-37, 39-46</p>

# RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Den Internationale No

PC1/FR 96/01244

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO-A-9428138	08-12-94	AU-A- 6760094	20-12-94
		CA-A- 2163641	08-12-94
		EP-A- 0700441	13-03-96
-----			